Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (Salmo salar)

Zueva, Ksenia J.

2018-06

Zueva , K J , Lumme , J , Veselov , A E , Kent , M P & Primmer , C R 2018 , ' Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (Salmo salar) ' , Marine Genomics , vol. 39 , pp. 26-38 . https://doi.org/10.1016/j.margen.2018.01.001

http://hdl.handle.net/10138/315575
https://doi.org/10.1016/j.margen.2018.01.001

Genomic signatures of parasite-driven natural selection in north European Atlantic salmon (*Salmo salar*).

Ksenia J. Zueva[1], Jaakko Lumme[2], Alexey E. Veselov[3], Matthew P. Kent[4], Craig R. Primmer[1,5,6]

[1]*Department of Biology, University of Turku, Turku, Finland ksezue@utu.fi*
[2]*Department of Biology, University of Oulu, Oulu, Finland*
[3] *Institute of Biology of the Karelian Research Centre, Petrozavodsk, Russia*
[4]*Centre for Integrative Genetics, Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, Ås, Norway*
[5]*Department of Biosciences, University of Helsinki, Helsinki, Finland*
[6]*Institute of Biotechnology, University of Helsinki, Helsinki, Finland*

ABSTRACT

Understanding the genomic basis of host-parasite adaptation is important for predicting the long-term viability of species and developing successful management practices. However, in wild populations, identifying specific signatures of parasite-driven selection often presents a challenge, as it is difficult to unravel the molecular signatures of selection driven by different, but correlated, environmental factors. Furthermore, separating parasite-mediated selection from similar signatures due to genetic drift and population history can also be difficult. Populations of Atlantic salmon (*Salmo salar L.*) from northern Europe have pronounced differences in their reactions to the parasitic flatworm *Gyrodactylus salaris* Malmberg 1957 and are therefore a good model to search for specific genomic regions underlying inter-population differences in pathogen response. We used a dense Atlantic salmon SNP array, along with extensive sampling of 43 salmon populations representing the two G. *salaris* response extremes (extreme susceptibility vs resistant), to screen the salmon genome for signatures of directional selection while attempting to separate the parasite effect from other factors. After combining the results from two independent genome scan analyses, 57 candidate genes potentially under positive selection were identified, out of which 50 were functionally annotated. This candidate gene set was shown to be functionally enriched for lymph node development, focal adhesion genes and anti-viral response, which suggests that the regulation of both innate and acquired immunity might be an important mechanism for salmon response to *G. salaris*. Overall, our results offer insights into the apparently complex genetic basis of pathogen susceptibility in salmon and highlight methodological challenges for separating the effects of various environmental factors.

Keywords: Atlantic salmon, genomic adaptation, genome scan, parasite-driven selection, *Gyrodactylus salaris*

INTRODUCTION

Parasites act as a strong selective force on natural populations, and given that many of them specialize on a single host species or a few host species, constant arms-races between the hosts and the parasite often occur (Carval and Ferriere, 2010; Kaltz and Shykoff, 1998). The mechanisms of host defence are very diverse and can focus on decreasing parasite fitness and increasing the host's ability to cope with the negative effects of the presence of a parasite (Carval and Ferriere, 2010; Råberg et al., 2007). Several genetic mechanisms of adaptive immune response in vertebrate hosts have been proposed, with earlier studies often focusing on the major histocompatibility complex (MHC) genes and occasionally focusing on other immune-relevant loci (Acevedo-Whitehouse and Cunningham, 2006; Medzhitov, 2007; Sommer, 2005). Being an important link to early vertebrate evolution, teleost fishes have also been the subject of intense research on the mechanisms of both innate and acquired immunity (Zhu et al., 2013). Special interest in understanding the basis of pathogen response had been given to salmonid fish species due to their commercial and recreational importance. For example, in Atlantic salmon, studies of the genetic basis of resistance have been conducted for a number of parasites and pathogens, including salmon lice (Holm et al., 2017; Tadiso et al., 2011), infectious pancreatic necrosis virus (Cepeda et al., 2011; Moen et al., 2015; Reyes-Lopez et al., 2015), anaemia virus (Moen et al., 2009), and the furunculosis-causing bacterium *Aeromonas salmonicida* (Dionne et al., 2009).

In this study, we concentrate on the genomic basis of Atlantic salmon adaptation to a particularly dangerous parasite, the monogenean flatworm *Gyrodactylus salaris*. Atlantic salmon from northern Europe exhibit striking differences in their susceptibility to the parasite. Salmon populations from rivers draining to the Atlantic Ocean and the Barents and White Seas are highly susceptible to *G. salaris*, with mortality rates following parasite exposure reaching 95% (Johnsen and Jensen, 1991). Landlocked populations from freshwater lakes Onega and Ladoga, however, are almost completely resistant, with low-level infections being observed in just 1% of fish (Kuusela et al., 2009). These differences are thought to be due to the phylogeographic histories of the regions, with land-locked salmon having a longer co-evolutionary history with the parasite in the eastern freshwater refugium (at least 130,000 years), whereas salmon from the Barents and White Seas were not exposed to *G. salaris* until recent decades (Kudersky et al., 2003; Kuusela et al., 2009, 2007; Lumme et al., 2016). Despite the potential threat to susceptible salmon populations following *G. salaris* introduction, the genetic basis of adaptation to *G. salaris* remains unclear, in spite of an increasing focus on this

topic (Gilbey et al., 2006; Kania et al., 2010; Matejusová et al., 2006; Tonteri et al., 2010; Zueva et al., 2014).

Understanding range-wide patterns of adaptation is challenging with experimental approaches, but approaches using genome-wide scans to detect signals of strong natural selection can provide means, albeit less direct, for identifying loci underlying local adaptation (Haasl and Payseur, 2016; Oleksyk et al., 2010). A strong selective force, which pathogen presence is likely to be (Fumagalli et al., 2011), is expected to result in an increase in the frequency of advantageous alleles, with a simultaneous reduction in variability in neutral linked sites (Nielsen et al., 2005; Oleksyk et al., 2010). In teleosts, genome scans and candidate gene approaches have been used to test whether genomic regions containing immune-relevant genes exhibit stronger evidence for selection compared to other regions (Pankratz et al., 2010; Tonteri et al., 2008), and to identify the genetic basis of local adaptation linked to a variety of other natural conditions, including water temperature and salinity (Guo et al., 2016; Kusakabe et al., 2017; Limborg et al., 2012b; Nielsen et al., 2009; Vilas et al., 2015). However, identifying genomic signals of selection in response to a particular selective pressure in wild populations can be challenging. One reason for this is that wild populations are constantly exposed to a plethora of different, and often correlated, selective pressures, the relative strengths of which are not always clear and/or may vary. Consequently, it can be challenging to predict which of the selective pressures leaves the most pronounced footprint in the genome and is thus the one that is most likely to be detected by natural selection scans. In addition, genomic signals of the effects of genetic drift (increased divergence and decreased diversity) can be similar to those of natural selection, albeit at a genome-wide scale. In populations with a small effective population size, the identification of signals of selection is even more difficult, as the effect of genetic drift on reduction in allele diversity is more pronounced (Schlötterer, 2003). Population history may further complicate the interpretation of selective signals, as even when subjected to the same selective pressure, populations with different phylogeographic histories, and therefore from different genetic lineages, may follow diverse adaptation paths due to dissimilarity in standing genetic variation (Przeworski et al., 2005). Therefore, a good understanding of population history can help minimize the number of false positives in scans for signals of natural selection.

The abovementioned challenges for identifying the genetic basis of adaptation are of relevance to contemporary Atlantic salmon from northern Europe. Following the retreat of the Scandinavian ice sheet after the last glacial maximum (17,000-15,000 years ago (ya)), different water basins have been colonized at different times and by salmon from various phylogenetic

lineages, resulting in the prolonged isolation of freshwater lakes Ladoga and Onega from Atlantic Ocean salmon and the pronounced genetic divergence between salmon populations in the region at various geographic scales. Northwest Russian lakes Onega and Ladoga were formed first, approximately 13,000 ya (Björck, 1995; Saarnisto and Saarinen, 2001), and were colonized by salmon from an eastern freshwater refugium, which had been previously isolated from an Atlantic Ocean influence for at least 130,000 years (Funder et al., 2002). The Kola Peninsula and White Sea areas were free of ice later than the Russian lakes and were re-colonized by salmon from refugia in the eastern Barents Sea and the south Atlantic Ocean (Asplund et al., 2004; Bourret et al., 2013; Tonteri et al., 2005). As a result, Baltic lineage salmon, including Onega and Ladoga stocks, are genetically highly diverged from the eastern Atlantic Ocean lineage that includes the Barents Sea and the White Sea (Asplund et al., 2004; Bourret et al., 2013; Nilsson et al., 2001; Ozerov et al., 2010; Tonteri et al., 2005). Furthermore, lower effective population sizes, and therefore an increased influence of genetic drift, have resulted in high divergence between the salmon populations from lakes Onega and Ladoga (Ozerov et al., 2010; Tonteri et al., 2007).

Given the prolonged isolation of freshwater salmon, it is likely that they have evolved a number of unique traits (in addition to *G. salaris* resistance) compared to populations in the rest of the range, including variation in the smoltification process (Kiiskinen et al., 2003; Nilsen et al., 2008, 2003) and other physiological functions (Peng et al., 2003) likely resulting from adaptation to a freshwater lifestyle. Water temperature profiles also differ between freshwater northwest Russian lakes and the northern Atlantic Ocean, with both the river water temperature during salmon development and the water temperature of salmon feeding grounds in lakes Ladoga and Onega being generally warmer (Naumenko et al., 1996; Tolstikov and Petrov, 2006). While temperature is known to affect metabolic and developmental rates (Brown et al., 2004; Gillooly et al., 2001), it also greatly influences food availability and trophic networks (Winder and Schindler, 2004) as well as pathogen diversity (Adlard et al., 2015; Dionne et al., 2007) and thus is likely to be a strong selective force both in lakes and the ocean. In addition, salmon populations in lakes Ladoga and Onega are relatively small in population size and therefore are likely to be characterized by strong genetic drift (Ozerov et al., 2010; Tonteri et al., 2007). Given the abovementioned, and given the fact that *G. salaris* susceptibility co-varies with gradients of salinity and temperature in North European salmon populations, it may be challenging to disentangle signals of parasite-mediated selection from other selective forces.

We have previously studied the genetic basis of Atlantic salmon adaptation to *G. salaris* using a genome scan based on 4,631 single nucleotide polymorphisms (SNPs) (Zueva et al.,

145 2014). To address the abovementioned challenges in identifying genetic footprints of selection,
146 we developed and implemented an analysis approach based on multiple tests for selection that
147 involves several combinations of populations varying in geographic location and susceptibility
148 to the parasite. Three genomic regions potentially involved in parasite resistance were
149 identified, as well as three regions possibly related to salinity adaptation. However, the limited
150 number of polymorphic markers, combined with the small number of population samples
151 available, maintains the possibility that some regions under selection may have been missed.
152 In the current study, we address those limitations by genotyping DNA pools on a 220,000 SNP
153 array, and increasing the number of surveyed populations from 12 to 43, allowing for a
154 considerable increase in the resolution of the selection signals.

MATERIALS AND METHODS

Ethics statement

Samples used in this study were obtained according to relevant national legislations and were described previously (Ozerov et al., 2012, 2010; Zueva et al., 2014).

Sampled populations

Atlantic salmon from 44 north European populations were initially included in the study. Sample sizes per location varied from 23 (Lizma river) to 326 (Tenojoki_1 river), totalling 2,438 individuals (Table 1, Figure 1). Most samples represented juveniles collected between 1997 and 2005 via electrofishing, where tissue was stored in 95% ethanol (see Ozerov et al. 2010, 2012 for details) except for the Näätämö River and two sub-populations from the Teno River that originated from air-dried scales collected from adults during their spawning migration (Aykanat et al., 2015; Pritchard et al., 2016). Earlier research has indicated that the vast majority of these population exhibit temporarily stable population structure (Ozerov et al., 2013).

Sample preparation and population pooling

Total genomic DNA was extracted using one of several methods including NucleoSpin® Tissue (Macherey Nagel) protocol, salt extraction protocol (Aljanabi and Martinez, 1997), vacuum extraction with glass beads (as in Elphinstone et al. 2003), or QIAamp DNA mini kit (Qiagen). DNA extraction and sample pooling for Tenojoki_1, Tenojoki_2 and Näätämö rivers were described in Aykanat et al. (2015) and Pritchard et al. (2016). Individual DNA samples from remaining 41 populations extracted for previous studies (e.g. Tonteri et al. 2007; Ozerov et al. 2010; Zueva et al. 2014), were subjectively assessed for degradation by electrophoretic separation on a 1% agarose gel. Samples showing excessive signs of degradation (low molecular weight DNA) were re-extracted with QIAamp DNA mini kit (Qiagen) (618 samples). The concentration of individual DNA samples was measured using a Qubit 2.0 fluorometer and Qubit dsDNA HS Assay kit (Life Technologies), and adjusted to a final concentration of 10 ng/ul. Equal amounts of DNA from all individuals from the same population were combined to make a population pool, with four technical replicates per population, i.e., 41 x 4 = 164 pools in total. The final concentration of each pool was measured with Qubit to verify that it was 10 ± 0.5 ng/ul.

SNP genotyping and data filtering

For each pool, allele intensities were obtained for 220,000 SNPs genotyped on a custom Affymetrix Axiom array (Life Technologies) according to the manufacturer's instructions (see Barson et al. 2015 for array details) at the Centre for Integrative Genetics (CIGENE, Norway). These data were subjected to a series of manipulations and quality checking steps. First, the relative intensities of the B allele were calculated and corrected for unequal allele representation using a polynomial specific probe correction algorithm, PPC (Brohede et al., 2005). Salmon (n=610) previously genotyped as individuals and allelotyped in pools were used for PPC correction (Supplementary material S_script1). SNP loci were removed from the data set if they, (i) did not include all 3 possible genotypes (AA, AB and BB) in the reference sample of 610 individuals, (ii) could be affected by a known off-target variant, (iii) deviated from HWE with $P<0.00001$, or (iv) had a minor allele frequency across all populations less than 0.05 (Supplementary material S_script2). SNPs were tested for deviation from Hardy-Weinberg equilibrium using individual genotypes of samples from the mainstem Tenojoki population (data as in Pritchard et al., 2016). Strong deviation from HWE may indicate genotype calling errors (e.g. homozygotes and heterozygotes are both being called as homozygotes), and a significance level of 0.00001 was chosen in order to primarily exclude SNPs that deviate from HWE due to technical issues. Furthermore, for every SNP, we tested the variability of allele B frequencies over four pooling replicates by comparing sets of SNPs with the highest standard deviation (SD) over replicas between all populations (20% of most variable SNPs for each population). None of the SNPs had high SD over replicates in all the populations, and therefore, none of the SNPs were filtered out during this step (Supplementary material S_script 3). Population Chapoma, however, was excluded due to a high number of SNPs with high SD over genotyping replicas (3% of all SNPs had SD > 0.1 in Chapoma, whereas for the other populations, this number was approximately 0.4%; Supplementary material S_script 3). For the remaining populations, the arithmetic mean of allele B frequencies for each SNP was calculated using allele frequencies falling within 25% and 75% quantiles of the original frequency distribution between four genotyping replicates (Supplementary material S_script 4, 5). After quality control, 197,431 SNPs and 43 populations were retained for further analyses.

Salmon genome annotation

Annotations for the Atlantic salmon genome were obtained via R package Ssa.RefSeq.db (https://github.com/FabianGrammes/Ssa.RefSeq.db). The package utilizes the latest publicly

available salmon genome build, ICSASG_v2 (https://www.ncbi.nlm.nih.gov/genome/369?genome_assembly_id=248466), and Gene Ontology annotations are assigned to genes by blasting the predicted coding sequences against the Swiss-Prot protein DB using Blast2GO software (Conesa et al., 2005) with the default settings. In cases when a gene had more than one transcript and thus more than one predicted protein sequence, the longest protein sequence was used in blastp. Mapping between the SNPs and the respective genes of interest was done using *bedtools* software (Quinlan and Hall, 2010), with SNP positions verified using NCBI dbSNP database (https://www.ncbi.nlm.nih.gov/projects/SNP/) and gene information obtained from the NCBI webpages of the ICSASG_v2 salmon genome build reference sequence files (https://www.ncbi.nlm.nih.gov/genome?LinkName=nuccore_genome&from_uid=925216783). Gene margins were defined as the region from the start of the 5´untranslated region (UTR) to the end of the 3' UTR, including the coding sequence for the longest predicted protein, and a SNP was assigned to a gene if its position in the genome fell within a gene margin.

Population genetics and outlier locus detection

Principal component analysis (PCA) was performed to assess the basic population genetic structure using the *prcomp* algorithm within the built-in '*stats*' package within the R-3.4.0 environment (R Core Team, 2016) (Supplementary material S_script 6). To identify the genome regions potentially affected by signals of selection, we used two independent methods that are suitable for implementation with pooled SNP data, and primarily concentrated on SNPs found to be outliers by both approaches.

Bayenv, implemented in the Bayenv2.0 software, is a Bayesian method that can be used to identify SNPs with unusually large allele frequency differentiation after accounting for population history and gene flow (Günther and Coop, 2013). As we were interested in selection signals that were common across populations within each geographical region, we treated the three geographic regions (Barents & White Seas, Lake Ladoga, and Lake Onega) as populations, while original populations within a region were treated as individuals. Allele frequencies per SNP per region were then calculated as the arithmetic mean of the population allele frequencies. Allele frequencies were converted to allele counts, required for Bayenv2.0 input files, based on the total number of individuals across all populations in a region (Supplementary material S_script7.1). Three pair-wise comparisons were performed: Barents & White Seas vs. Ladoga lake, Barents & White Seas vs. Onega lake, and Ladoga lake vs.

Onega lake. Using PLINK software (Purcell et al., 2007) we first identified SNPs with no linkage disequilibrium between them, and a random subset of these SNPs was used to compute a covariance matrix between the populations. We checked the convergence of matrices built using different numbers of random SNPs and the influence of the iteration number. As a result, covariance matrices were built using 50,000 SNPs and 100,000 iterations, which proved to be both sufficient and computationally effective. During the next step, Bayenv2.0 calculates a population differentiation statistic called X^TX, analogous to the well-known F_{ST}, but based on standardized allele frequencies that were derived to account for population structure. X^TX was calculated using 10,000 iterations and was used to identify loci that are more differentiated than expected under pure drift between populations (Günther and Coop, 2013). The software does not provide significance estimations for deviation from the null distribution. Therefore, a custom cut-off at the upper 0.005% quantile of the statistical distribution was applied to determine possible SNP outliers (Supplementary material S_script7.2). Next, we compared the results from the three pair-wise comparisons, focusing on genomic regions harbouring peaks of SNPs with elevated X^TX statistics in both the Barents & White Seas vs. Ladoga lake and the Barents & White Seas vs. Onega lake comparisons, but absent from the Ladoga lake vs. Onega lake comparison. Populations in landlocked lakes have been isolated from each other for a long period of time, and this approach allows us to exclude genomic regions that are likely to exhibit elevated levels of differentiation due to genetic drift rather than directional selection. Further, we identified genes that contained outlier SNPs within their margine using *bedtools* software and the procedure described above. Only genes that had outliers in both Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake tests were considered to be candidates (see Results).

A second method used to detect outlier loci was the Bayesian approach implemented in the BayeScan2.1 software (Foll and Gaggiotti, 2008). This approach allows direct estimation of the posterior probability of a given locus to be under the effect of selection by defining two alternative models, one that includes the effect of selection and another that excludes it, and testing their respective posterior probabilities using a MCMC approach. The method uses population-specific and locus-specific components of F_{ST} coefficients and has been suggested to be robust when dealing with complex demographic scenarios for neutral genetic differentiation (Foll and Gaggiotti, 2008). The same logic as for *Bayenv2* was applied: we performed three pair-wise comparisons, where geographic regions were treated as populations and original populations were treated as individuals. Calculations were performed under the default parameters. Outlier SNPs were identified with a false discovery rate of 0.05, and

respective genes were assigned to each outlier SNP using *bedtools*. The final set of candidate genes under selection was obtained by identifying genes that were detected by both Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake comparisons, but were not present among outliers in the Ladoga lake vs. Onega lake test. Genes present both in Bayenv2.0 and BayeScan2.1-based candidate gene sets were further considered as candidate genes potentially affected by selection (Supplementary material S_script8).

We also considered single SNPs with the most pronounced levels of genetic differentiation based on both Bayenv2.0 and BayeScan2.1 approaches as candidates, initially regardless of the overlap between the tests or population comparisons. For each of the tests, the 50 SNPs with the most extreme statistics (the highest X^TX for Bayenv2.0, and the lowest q-values for BayeScan2.1) were selected, and associated genes harbouring the SNPs were retrieved using the *bedtools*. These gene sets were then compared between the Bayenv2.0 and BayeScan2.1 approaches and between the Barents & White Seas vs. Ladoga lake, Barents & White Seas vs. Onega lake and Ladoga lake vs. Onega lake pair-wise comparisons.

Functional annotation and functional enrichment test

Annotations of all SNPs to specific gene ontology (GO) terms (Ashburner et al., 2000) were obtained via the Ssa.RefSeq.db package. To determine whether the set of identified candidate genes (see Results) was significantly enriched or depleted for particular GO terms, we performed an enrichment test, implemented in the *topGO* package in R, using the *weight01* algorithm and the list of all Atlantic salmon genes that contained SNPs from the SNP array as a reference (Supplementary material S_script9).

RESULTS

Population genetic structure

Populations from different geographic regions clustered in distinct groups based on principal component analysis. The first PC explained 42% of the variance and separated the Barents & White Seas salmon populations from the freshwater lakes. The second component, explaining 9% of variance, separated the Ladoga and Onega lakes (Figure 2).

Detecting signals of selection

X^TX statistics for each SNP for each of the Barents & White Seas vs. Ladoga lake, Barents & White Seas vs. Onega lake and Ladoga lake vs. Onega lake pair-wise tests and 0.005% upper quantile outliers were estimated using *Bayenv2* and are presented in Table_S1. Altogether, 118 candidate outlier genes were detected in both Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake comparisons, but not from the Ladoga lake vs. Onega lake test after mapping outlier SNPs to specific genes (Figure 3, Table_S2).

Outlier SNPs detected using *BayeScan2.1* for each of the Barents & White Seas vs. Ladoga lake, Barents & White Seas vs. Onega lake and Ladoga lake vs. Onega lake pair-wise tests with a q-value false discovery threshold of 0.05 are presented in Table_S3. Once SNPs were mapped to the genes, 167 candidate genes harbouring SNPs exhibiting signals consistent with positive selection were shared between the Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake tests (Figure 3, Table_S4).

Fifty seven genes with outlier SNPs on 24 chromosomes were common for both the Bayenv2 and BayeScan2.1 analysis approaches and were therefore considered as the most promising candidates to be affected by positive selection that distinguishes salmon originating from the Atlantic Ocean from those originating from freshwater lakes (Table 2, Table_S5).

When the most highly genetically differentiated SNPs from both Bayenv2.0 and BayeScan2.1 tests were considered, approximately half of the yielded genes were similar between both approaches (Table_S6). However, the overlap between the Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake comparisons was less pronounced. Only one gene, serine/threonine-protein phosphatase regulatory subunit, was among the most differentiated based on both Bayenv2.0 and BayeScan2.1 and for both the Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake comparisons (Table 2).

Functional annotation

342 Using the Ssa.RefSeq.db package and the annotation procedure described above, out of
343 the 48,785 protein-coding genes reported for Atlantic salmon, 30,560 genes had a SNP within
344 their margin, and 23,850 of these genes were annotated with GO terms (Table_S7). From a
345 candidate set of 57 genes, 50 genes were annotated (Table_S7). Enrichment analysis retrieved
346 several GO terms that were significantly overrepresented in the set of 50 annotated candidate
347 genes. These included three biological processes with significance levels less than 0.01:
348 response to arsenic containing substance (GO: 0046685), lymph node development (GO:00
349 48535), and response to virus (GO: 0009615). The only enriched cellular components was
350 microtubule organizing center (GO:0005815). Finally, the most highly enriched molecular
351 function GO terms were phospholipase activator activity (GO: 0016004), vinculin activity
352 (GO: 0017166), and phosphatidylinositol phospholipase C activity (GO: 0004435) (Table_S5,
353 S8).

DISCUSSION

In this study, we utilized an extensive sample of 43 anadromous and landlocked salmon populations and almost 200,000 SNPs to investigate the genomic basis of differences in susceptibility to the parasite *Gyrodactylus salaris* observed in north European Atlantic salmon populations. By combining results from different outlier tests, we established a set of 57 candidate genes potentially associated with parasite tolerance/susceptibility.

Evidence of immune related functions of candidate genes

Several GO terms related to both innate and acquired immunity were enriched amongst the 57 candidate genes.

The most significantly enriched biological process term was response to arsenic-containing substance, GO:0046685, and along with another significant term, response to virus (GO:0009615), it was associated with the interferon-induced GTP-binding protein Mx-like gene. There are several copies of the mx (myxovirus)-like gene on chromosome 25, and three of them are included in the set of 57 candidate genes in our analysis (Table 2). Mx genes are induced during virus infection as a part of interferon-mediated innate immune response, and are active against a wide range of DNA and RNA viruses (Mitchell et al., 2013). The same set of genes was recently found to be highly diverged between salmon lineages from the Atlantic Ocean and the Finnmark region, which includes the Barents Sea clade (Kjærner-Semb et al., 2016). Our results suggest this genome region is under selection in Atlantic salmon more broadly as our study included populations from regions not studied previously, e.g. land-locked populations.

The third significantly enriched biological process term was lymph node development, GO:0048535. Earlier research on a number of genes with this GO term provides support for this process potentially being important in *Gyrodactylus* resistance in Atlantic salmon. Lymph nodes are essential part of the mammalian adaptive immune system as they are involved in lymph filtering and circulation, and are a place of residence for leukocytes including B and T lymphocytes. The lymphoid system of teleosts lacks lymph nodes, but include organs with similar functions such as lymphatic vessels, thymus, head kidney (considered an ortholog of mammalian bone marrow) and spleen (Hedrick et al., 2013; Sunyer, 2013). T-lymphocytes, developing and maturing in the thymus, as well as interleukin signalling pathways in fishes also resemble those of mammals (Nakanishi et al., 2015; Zapata et al., 2006). One of the genes associated with the lymph node GO term is the T-cell leukemia homeobox protein 1 (*TLX1*) -

like gene, involved in mammalian spleen development (Yamamoto et al., 1995). *TLX1* is also expressed during fish spleen development and presumably contributes to a supportive microenvironment for the maturation of lymphocytes, which appear in fish spleen after they become present in thymus and kidney (Boehm et al., 2012; Li et al., 2017). Another associated gene, nuclear receptor ROR-alpha-like gene (*RORα)*, has diverse biological functions including regulation of glucose and free fatty acid metabolism (Kadiri et al., 2015) and is also an important pro-inflammatory agent participating in regulation of inflammation cytokines (Sadeghi et al., 2015; Sun et al., 2015) and modulation essential for inflammation T-helper lymphocytes (Th-17) (Yang et al., 2008). In teleosts, *RORα*, along with another transcription factor *RORγ*, regulate expression of pro-inflammatory interleukins-17 (IL-17). IL-17 members have been identified in several fish species including Atlantic salmon, and are reported to play crucial roles in host defense against microbial organisms (Chi et al., 2016; Kumari et al., 2009).

Involvement of the candidate gene set in immune response was further emphasized by the enriched cellular component and molecular function GO terms: microtubule organizing center, vinculin binding, and phospholipase activity; united by their association to cytoskeleton, formation of focal and cell-cell adhesions and cell signaling. Reorganization of leukocytes' membrane, required for generating and maintaining immune response, depends on drastic changes of microtubule organizing center and involves the segregation of membrane and intracellular signaling proteins (Sancho et al., 2002). Talin, one of the proteins associated with the microtubule organizing center (MOC) GO term (Table_S5), binds trans-membrane receptors to actin cell cytoskeleton and is crucial during phagocytosis in amoeba and mammals (Freeman and Grinstein, 2014; Lim et al., 2007), as well as during adhesion of natural killer cells and T-lymphocytes to the extra-cellular matrix and target cells (Mace et al., 2009; Stanton et al., 2014). Talin-integrin complex is stabilized with the help of vinculin, and is dependent on activating of tyrosine phosporylation, as well as activity of phosphatidylinositol 3-kinase (PI3K) and phospholipase C. Altogether two genes from the candidate gene set have phospholipase A and phospholipase C- activity. Phospholipases, while involved in a number of signalling pathways, are known to have an important role in signal transduction in leukocytes, e.g., natural killer cells (Caraux et al., 2017), and have a pronounced role in inflammation processes (Boilard et al., 2010). Most of the abovementioned studies were conducted using mammalian models, but teleosts express a variety of similar interleukins, integrin complexes, and signalling pathways involved in immune regulation, for example IL-21,22 and PI3K-signalling (Costa et al., 2013; Wang et al., 2011), immunoreceptors that contain tyrosine-based activation motifs (ITAMs) (Blank et al., 2009), as well as CR3 integrins

that require actin-activation (Lillico et al., 2017; Mikrou et al., 2009). Toll-signaling, another vital part of immune signaling, is also present in teleost fishes (Hanington et al., 2009; Rebl et al., 2010). One of the candidate genes, sphingomyelin phosphodiesterase 3, is involved in lipid metabolism and was shown to regulate Toll-like receptors signalling in mice macrophages (Heinz et al., 2015). Granulin, one of the co-factors for Toll-like receptors (Park et al., 2011), promotes host cell proliferation when excreted by liver flukes (Bansal et al., 2017; Smout et al., 2009), and it is known that an extensive skin and mucus proliferation is one of the consequences of *G. salaris* infection in susceptible Atlantic salmon (see below).

Three of the candidate genes harbour non-synonymous (missense) outlier SNPs, which implies that they may result in a change of protein structure and therefore possibly protein function (Table2, Table S5). Linking these genes to immune response processes is not straightforward, as they are involved in a number of cellular processes. However, *NAGPA*, playing part in lysosomal activity, is known to be important in maturation of dendritic cells required for T-cells stimulation (Trombetta et al., 2003). RNA helicases from the DEAD/H family, to which another gene, DNA helicase *ddx11*, belongs to, have been associated with innate immunity and response to viruses in humans (Oshiumi et al., 2010; Schröder, 2011) and salmonids (Castro et al., 2013; Krasnov et al., 2011). A link with pathogen-induced signalling in innate immune system was also identifiable for the third gene, *TBC1D5*, which is involved in induction and regulation of autophagy (Faure and Lafont, 2013). Taken together, these results suggest involvement of the candidate gene set in cell-signalling during both innate and adaptive immune response, and the mentioned genes are thus promising candidates for future research.

The candidate gene set described above was formed based on criteria of overlap between the Bayenv2.0 and BayeScan2.1 results and overlap between the Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake pair-wise comparisons. However, we also looked for the genes harbouring the most significant SNPs regardless of the overlaps between the analysis approaches and pair-wise comparisons. The Bayenv2.0 and BayeScan2.1 tests independently resulted in comparable sets of SNPs with high differentiation, with approximately 50% of the related genes being similar between the tests. Within the results of both Bayenv2.0 and BayeScan2.1 tests, there were obvious differences between the pair-wise comparisons, as almost none of the genes overlapped between the Barents & White Seas vs. Ladoga lake, Barents & White Seas vs. Onega lake, or Ladoga lake vs. Onega lake pair-wise comparisons. Only one gene from the candidate gene set described above, the serine/threonine-protein phosphatase 2A (*pp2A*) 56 kDa regulatory subunit gamma -like gene, emerged in both

the Barents & White Seas vs. Ladoga lake and Barents & White Seas vs. Onega lake comparisons (but not in the Ladoga lake vs. Onega lake comparison). PP2A is the major phosphatase for microtubule-associated proteins (Abraham et al., 2000; Watkins et al., 2012) and is abundant in mammalian lymphocytes (Chuang et al., 2000). It was suggested as an immune relevant gene in common carp (Abdelkhalek et al., 2014), and is thus a promising gene for future research on immune regulation in fishes.

Overall, these findings are consistent with the idea of *G. salaris* tolerance/susceptibility differences observed between landlocked and Atlantic Ocean salmon populations being linked to natural selection acting on the regulatory mechanisms of both innate and adaptive immune systems. To further test this hypothesis and to strengthen the candidacy of reported genes several directions for future research could be suggested. First, the expression profiles for the candidate genes can be compared by parasite challenge experiments using individuals from susceptible and resistant populations in a controlled laboratory environment, e.g. using qPCR. This approach was fruitful in gene expression experiments featuring Atlantic salmon and *G. salaris*, when up-regulation of potentially immune relevant myeloid leukemia differentiation protein was detected in susceptible salmon (Matejusová et al., 2006); as well as in number of other studies challenging salmonids with various pathogens (Haarder et al., 2013; Krasnov et al., 2012). Furthermore, for candidate genes with presumed regulatory function, expression of downstream regulatory targets could also be quantified, and for candidate genes with presumed enzymatic activity (effectors) biochemical assays could be used to quantify corresponding protein activity. Another possibility for follow-up research is to concentrate on *G. salaris* susceptible salmon populations that due to continuous re-stocking survive the infection: Keret' river in the White Sea (Kuusela et al., 2009), and Drammen system in Norway (Bakke et al., 1990). Notably, genetic composition of Keret' river salmon was suggested to have changed over the years in response to parasite load (Artamonova et al., 2008). Utilizing a dense SNP array and a knowledge of candidate genes, it is possible to test hypotheses of the precise targets of this temporal change.

Comparison with previous studies

One of the genes from the candidate genes set, coding for an adhesion G protein-coupled receptor L2 (*adgrl2*), is located within the *G. salaris* - related region on chromosome 10 detected in our previous study (Zueva et al., 2014). G protein-coupled receptors are involved in a plethora of signal-transduction pathways, including T-cell signaling (Goetzl et al., 2004; Smit et al., 2007), and are common among various taxa (Schiöth and Fredriksson, 2005).

ADGRL-type receptors, at least in mammals, are crucial for functioning of nervous and cardiovascular systems, but their possible involvement in teleost immune system is not yet clear.

Taken together, genomic regions associated with *G. salaris* presence based on Zueva et al. (2014) were clustered into two functional groups; one of the groups included genes involved in T cell activation and the other included genes involved in the synthesis and elongation of fatty acids, which are known to moderate inflammation and act as anti-pathogen agents (Calder, 2001; Carballeira, 2008; Harbige, 2003). Regulation of T-lymphocytes is part of adaptive immunity, whereas lipid metabolism and macrophage activation are part of innate immunity. Experimental studies on Atlantic salmon response to *G. salaris* are limited, and it is hard to predict the exact mechanisms that form the foundation of resistance and/or tolerance to the parasite. It was demonstrated, however, that highly susceptible salmon from the east Atlantic responded to *G. salaris* exposure by an elevated production of interleukin-1b and interferon-gamma cytokines, which enhance the proliferation of the epithelial and mucous cells that the parasite feeds on. Less susceptible Baltic salmon responded to the parasite with delay and by the activation of genes that did not result in mucus proliferation. It was suggested that by regulating the initial stages of inflammation, and consequently, mucus production, Baltic salmon are able to control parasite abundance by starving it (Kania et al., 2010; Lindenstrom et al., 2006). An acquired immune response generally takes more time to develop and initiate, and given that *G. salaris* presence can result in the rapid decline of infected fish (Bakke et al., 1990, 2004), it is feasible that the defence mechanisms against this parasite species are focused, at least partly, around the innate immune system.

The overall functional patterns of the detected gene sets are similar between the previous and current studies, and the limited overlap between the candidate genes and genomic regions is not unexpected. Most obviously, the current study is based on a qualitatively greater number of SNPs, with the average SNP density being one SNP per 0.018 mega bases (Mb) as opposed to one SNP per 0.5 Mb in the previous study. Consequently, previous regions of elevated F_{ST} are masked by much more narrow and abundant regions of both elevated and reduced genetic differentiation that have been detected in current study. The low SNP coverage in Zueva et al. 2014 also complicates the result comparison itself, as whether we find overlap or not depends on the distance used to assign SNPs to genes. In addition, the annotation of the salmon genome has improved rapidly in recent years following publication of the Atlantic salmon genome sequence (Lien et al., 2016). Indeed, a number of significant SNPs were excluded from enrichment analyses in the previous study due to a lack of functional annotation. Furthermore,

the two studies differ in the methods used for identifying genomic regions of selection and in the analysis design used to select the final candidate gene sets. In addition, while there was only one overlap between the current results and the genomic regions identified based on a combination of all four tests in the previous study, there were common genes when considering the tests used in Zueva et al. (2014) one at a time. For example, Design 4 (single loci outlier test) from Zueva et al. (2014) detected one gene that was also present in the current candidate gene set: *wwc1*, participating in phosphorylation; Design 1 (reduced diversity in freshwater lakes) detected a *RORα*-like gene described previously, and a chromobox protein homolog 7 that among other functions modulated $CD4^+$ T cell apoptosis in mammals (Li et al., 2014). Overall, it is encouraging that regardless of the chosen strategy, both the current and previous studies resulted in identifying candidate regions that share functional characteristics.

The genomic basis of tolerance to *G. salaris* has also been studied using a QTL approach by back-crossing the parasite susceptible Scottish salmon with parasite-tolerant Baltic salmon, and several microsatellites associated with *G. salaris* tolerance have been identified (Gilbey et al., 2006). These associations represented entire linkage groups, and a direct comparison of the results should be done with caution since there can be inconsistencies in linkage group names between the SALMAP project, used by Gilbey and co-authors, and the current Atlantic salmon genome build. Nevertheless, linkage groups 1, 4, 5, 6, 9, 13, 18 and 25 were suggested by both results, and altogether, our findings are consistent with the idea of polygenic control for both innate and acquired *G. salaris* resistance as suggested by Gilbey and co-authors.

2. Interpreting the results: challenges and perspectives

Biological perspective: the challenge of correlated environmental traits.

The results of this study highlight several of the challenges of using a genome-scan approach to identify loci associated with a specific phenotypic trait, even when dramatic differences in the trait exist between replicated populations. These challenges can be both environmental and genetic in nature. As noted earlier, separating signals of selection on correlated environmental and phenotypic traits can be challenging when working at the between-population level. In the case of Atlantic salmon from northern Europe, these traits include parasite presence/absence, salinity of the water basin that the fish migrates to, water temperature in both the home rivers and on the feeding grounds, as well as hypothetical differences in fish diet in marine and freshwater environments. Given the potential drastic effect *G. salaris* has on fish survival, our assumption was that parasite presence should leave a

very pronounced footprint of selection in the salmon genome, and our analyses were designed to focus on this assumption. However, we cannot exclude the possibility that the observed signals of selection are partly due to other selective forces in addition to the effects of the parasite. For example, apart from immune function, actin-based cytoskeleton was shown to play a role in osmotic regulation of $K^+/Na^+/2Cl$ cotransporters (Flatman, 2002; Lionetto and Schettino, 2006). Cation-chloride cotransporters, such as the $K^+/Na^+/2Cl$ cotransporter, are known to be associated with salinity adaptation in a number of fish species, including the brackish medaka (*Oryzias dancena*) (Kang et al., 2010), Mozambique tilapia, *Oreochromis mossambicus*, (Hiroi et al., 2008) and mummichog, *Fundulus heteroclitus* (Hoffmann et al., 2002). Interleukins, immune signal molecules associated with a number of genes desribed above, have also shown signals of divergent selection between anadromous and landlocked brown trout (Limborg et al., 2012a; Narum et al., 2011). On the other hand, a number of studies have documented an increase of phagosystosis, alterations in antimicrobial enzyme lysozyme levels, as well as change in IgM levels in response to salinity alteration, indicating strong effect of salinity on innate and adaptive immune systems of teleosts (Bowden, 2008; Makrinos and Bowden, 2016). A genetic issue that may disguise the target of selection is gene pleiotropy: when a particular gene has multiple functions, it may not be clear which specific function has resulted in a gene or a genomic region exhibiting signatures of selection. Furthermore, pleiotropic effects may constrain selection on a particular trait, when the genetic response to selection on one trait is limited by selection on other correlated traits controlled by the gene (Orr, 2000; Wagner and Zhang, 2011). The level of gene pleiotropy has been shown to be negatively correlated with variability in gene expression in response to environmental change and is thus an evolutionary constraint (Papakostas et al., 2014). Many candidate genes from our study are involved in a number of other processes apart from immunity. For example, serine/threonine-protein phosphatase 2A and 2B are also involved in osmoregulation (Nakamura et al., 1993; Shiozaki and Russell, 1995), while nuclear receptor ror-alpha is associated with circadian clock (Yang et al., 2006). If the basis of salmon response to *G. salaris* is controlled by several genes with multiple additional functions, the genomic signals of selection on these genes may be less pronounced and thus more difficult to identify and/or interpret.

Methodological perspective: power of genome scans

The candidate gene set identified in this study is based on the overlapping results of two approaches for identifying signals of selection, implemented in the Bayenv2.0 and

BayeScan2.1 software. A focus on loci identified as outliers in several tests has been applied in a number of studies to strengthen the candidacy of identified loci targeted by selection and to reduce type I errors (Oleksyk et al., 2008; Vasemägi et al., 2005). Such an approach may, however, reduce the chances of identifying loci under weak selection (Whitlock and Lotterhos, 2015). Both approaches implemented in the current study suggested a large number of "outlier" SNPs and associated genes, but just under half were common between the tests (57 genes out of more than 115 in each test). These tests are based on quantifying population differentiation in terms of F_{ST} or related measures and use different approaches to correct for neutral population structure (Hoban et al., 2016); thus, some of the resulting outliers might be an outcome of pronounced genetic drift and restricted gene flow (Bierne et al., 2011; Oleksyk et al., 2010), explaining the lack of full overlap between the identified regions of selection.

One caveat that is relevant for our study, and indeed all outlier analyses conducted in Atlantic salmon thus far is that due to not complete genome rediploidization approximately 10% of Atlantic salmon genome retain residual tetrasomy (Lien et al., 2016). Because of this, SNPs from this portion of the genome are not represented in the SNP array for technical reasons. Therefore potentially important genes residing in those regions might not have been detected. Another factor that could potentially affect outlier identification is ascertainment bias (Lachance and Tishkoff 2013) stemming from the fact that SNPs included in the array were based on their polymorphism in Norwegian aquaculture salmon from the Atlantic lineage. However, the relative levels of population genetic diversity and divergence estimated in the present study are in line with previous assessments using a different marker type (microsatellites: Ozerov et al., 2010; Tonteri et al., 2009). Further, all comparisons are between multiple populations from lineages other than the Atlantic lineage. Thus, it is unlikely that potential SNP ascertainment bias has had a large effect on the results.

CONCLUSIONS

Overall, our results suggest an apparently complex genetic basis of *Gyrodactylus salaris* susceptibility and resistance in Atlantic salmon and highlight some methodological challenges for separating the effects of various environmental factors. Despite these challenges, it appears that the regulation of both innate and acquired immunity are important mechanisms in the response of Atlantic salmon to *G. salaris* and this study provides a number of promising candidate genes for future studies.

ACKNOWLEDGMENTS

We would like to express our gratitude to Jan Nilsson, Jaakko Erkinaro, Anti Vasemägi, and others involved in field collecting of the samples used in the study. Great thanks to

Mikhail Ozerov and Anni Tonteri for performing some of the DNA extractions more than 10 years ago. We are grateful to Fabian Grammes for his help with the Ssa.RefSeq.db package. Special thanks to Victoria Pritchard and Matthieu Bruneaux for analytical assistance and to the reviewers for their constructive comments that improved the manuscript. This research has been funded by the Academy of Finland (grants 302873 and 284941).

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: KJZ, CRP. Performed the experiments: KJZ, MPK. Analyzed the data: KJZ. Contributed reagents/materials/analysis tools: AEV, JL, MPK. Wrote the paper: KJZ, CRP. Commented on the manuscript: MPK, JL, AEV.

DATA ARCHIVING STATEMENT

Raw data and code used in analyses are archived in the Dryad Digital Repository: (http://dx.doi.org/..., to be updated)

REFERENCES

Abdelkhalek, N.K., Komiya, A., Kato-Unoki, Y., Spmamoto, T., Nakao, M., 2014. Identification of Immune Relevant Genes Using Expressed Sequence Tags (ESTs) in Common carp (Caprinus caprio) Gills. J. Fac. Agric. Kyushu Univ. 59, 109–116.

Abraham, D., Podar, K., Pacher, M., Kubicek, M., Welzel, N., Hemmings, B.A., Dilworth, S.M., Mischak, H., Kolch, W., Baccarini, M., 2000. Raf-1-associated protein phosphatase 2A as a positive regulator of kinase activation. J. Biol. Chem. 275, 22300–22304. doi:10.1074/jbc.M003259200

Acevedo-Whitehouse, K., Cunningham, A.A., 2006. Is MHC enough for understanding wildlife immunogenetics? TRENDS Ecol. Evol. 21, 433–438.

Adlard, R.D., Miller, T.L., Smit, N.J., 2015. The butterfly effect: parasite diversity, environment, and emerging disease in aquatic wildlife. Trends Parasitol. 31, 160–166. doi:10.1016/j.pt.2014.11.001

Aljanabi, S.M., Martinez, I., 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acids Res. 25, 4692–4693.

Artamonova, V.S., Khaimina, O. V., Makhrov, A.A., Shirokov, V.A., Shulman, B.S., Shurov, I.L., 2008. Evolutionary consequences of parasite invasion: A case study of Atlantic salmon Salmo salar L. Dokl. Biol. Sci. 423, 412–415. doi:10.1134/S0012496608060136

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. Nat. Genet. 25, 25–29. doi:10.1038/75556

Asplund, T., Veselov, A., Primmer, C.R., Bakhmet, I., Potutkin, A., Titov, S., Zubchenko, A., Studenov, I., Kaluzchin, S., Lumme, J., 2004. Geographical structure and postglacial history of mtDNA haplotype variation in Atlantic salmon (Salmo salar L.) among rivers of the White and Barents Sea basins. Ann. Zool. Fennici 41, 465–475.

Aykanat, T., Johnston, S.E., Orell, P., Niemelä, E., Erkinaro, J., Primmer, C.R., 2015. Low but significant genetic differentiation underlies biologically meaningful phenotypic divergence in a large Atlantic salmon population. Mol. Ecol. 24, 5158–5174. doi:10.1111/mec.13383

Bakke, T.A., Jansen, P.A., Hansen, L.P., 1990. Differences in the host resistance of Atlantic salmon . Salmo salar L ., stocks to the monogenean Gyrodactylus salaris Malmberg , 1957. J. Fish Biol. 37, 577–587.

Bakke, T. a, Harris, P.D., Hansen, H., Cable, J., Hansen, L.P., 2004. Susceptibility of Baltic and East Atlantic salmon Salmo salar stocks to Gyrodactylus salaris (Monogenea). Dis. Aquat. Organ. 58, 171–7. doi:10.3354/dao058171

Bansal, P.S., Smout, M.J., Wilson, D., Cobos Caceres, C., Dastpeyman, M., Sotillo, J., Seifert, J., Brindley, P.J., Loukas, A., Daly, N.L., 2017. Development of a Potent Wound Healing Agent Based on the Liver Fluke Granulin Structural Fold. J. Med. Chem. 60, 4258–4266. doi:10.1021/acs.jmedchem.7b00047

Barson, N.J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G.H., Fiske, P., Jacq, C., Jensen, A.J., Johnston, S.E., Karlsson, S., Kent, M., Moen, T., Niemelä, E., Nome, T., Næsje, T.F., Orell, P., Romakkaniemi, A., Sægrov, H., Urdal, K., Erkinaro, J., Lien, S., Primmer, C.R., 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. Nature 528, 405–408. doi:10.1038/nature16062

Bierne, N., Welch, J., Loire, E., Bonhomme, F., David, P., 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. Mol. Ecol. 20, 2044–2072. doi:10.1111/j.1365-294X.2011.05080.x

Björck, S., 1995. A Review of the history of the Baltic Sea, 13.0-8.0 ka BP. Quat. Iternational

691 27, 19–40.
692 Blank, U., Launay, P., Benhamou, M., Monteiro, R.C., 2009. Inhibitory ITAMs as novel
693 regulators of immunity. Immunol. Rev. 232, 59–71. doi:10.1111/j.1600-
694 065X.2009.00832.x
695 Boehm, T., Hess, I., Swann, J.B., 2012. Evolution of lymphoid tissues. Trends Immunol. 33,
696 315–321. doi:10.1016/j.it.2012.02.005
697 Boilard, E., Lai, Y., Larabee, K., Balestrieri, B., Ghomashchi, F., Fujioka, D., Gobezie, R.,
698 Coblyn, J.S., Weinblatt, M.E., Massarotti, E.M., Thornhill, T.S., Divangahi, M.,
699 Remold, H., Lambeau, G., Gelb, M.H., Arm, J.P., Lee, D.M., 2010. A novel anti-
700 inflammatory role for secretory phospholipase A2 in immune complex-mediated
701 arthritis. EMBO Mol. Med. 2, 172–187. doi:10.1002/emmm.201000072
702 Bourret, V., Kent, M.P., Primmer, C.R., Vasemägi, A., Karlsson, S., Hindar, K., McGinnity,
703 P., Verspoor, E., Bernatchez, L., Lien, S., 2013. SNP-array reveals genome-wide
704 patterns of geographical and potential adaptive divergence across the natural range of
705 Atlantic salmon (Salmo salar). Mol. Ecol. 22, 532–551. doi:10.1111/mec.12003
706 Bowden, T.J., 2008. Modulation of the immune system of fish by their environment. Fish
707 Shellfish Immunol. 25, 373–383. doi:10.1016/j.fsi.2008.03.017
708 Brohede, J., Dunne, R., McKay, J.D., Hannan, G.N., 2005. PPC: An algorithm for accurate
709 estimation of SNP allele frequencies in small equimolar pools of DNA using data from
710 high density microarrays. Nucleic Acids Res. 33, 1–7. doi:10.1093/nar/gni142
711 Brown, Ja.H., Gillooly, J.F., Allen, A.P., Savage, V.M., West, G.B., 2004. Toward a
712 metabolic theory of ecology. Ecology 85, 1771–1789.
713 Calder, P.C., 2001. Polyunsaturated fatty acids, inflammation, and immunity. Lipids 36,
714 1007–1024.
715 Caraux, A., Kim, N., Bell, S.E., Zompi, S., Ranson, T., Lesjean-Pottier, S., Garcia-Ojeda,
716 M.E., Turner, M., Colucci, F., 2017. Phospholipase C- y2 is essential for NK cell
717 cytotoxicity and innate immunity to malignant and virally infected cells. Blood 107,
718 994–1003. doi:10.1182/blood-2005-06-2428.Supported
719 Carballeira, N.M., 2008. New advances in fatty acids as antimalarial, antimycobacterial and
720 antifungal agents. Prog. Lipid Res. 47, 50–61. doi:10.1016/j.plipres.2007.10.002
721 Carval, D., Ferriere, R., 2010. A unified model for the coevolution of resistance, tolerance,
722 and virulence. Evolution (N. Y). 64, 2988–3009. doi:10.1111/j.1558-5646.2010.01035.x
723 Castro, V., Grisdale-Helland, B., Jørgensen, S.M., Helgerud, J., Claireaux, G., Farrell, A.P.,
724 Krasnov, A., Helland, S.J., Takle, H., 2013. Disease resistance is related to inherent
725 swimming performance in Atlantic salmon. BMC Physiol. 13, 1–12. doi:10.1186/1472-
726 6793-13-1
727 Cepeda, V., Cofre, C., González, R., Mackenzie, S., Vidal, R., 2011. Identification of genes
728 involved in immune response of Atlantic salmon (Salmo salar) to IPN virus infection ,
729 using expressed sequence tag (EST) analysis. Aquaculture 318, 54–60.
730 doi:10.1016/j.aquaculture.2011.04.045
731 Chi, H., Bøgwald, J., Dalmo, R.A., Zhang, W., Hu, Y. hua, 2016. Th17 master transcription
732 factors RORα and RORγ regulate the expression of IL-17C, IL-17D and IL-17F in
733 Cynoglossus semilaevis. Dev. Comp. Immunol. 55, 169–178.
734 doi:10.1016/j.dci.2015.11.001
735 Chuang, E., Fisher, T.S., Morgan, R.W., Robbins, M.D., Duerr, J.M., Vander Heiden, M.G.,
736 Gardner, J.P., Hambor, J.E., Neveu, M.J., Thompson, C.B., 2000. The CD28 and CTLA-
737 4 Receptors Associate with the Serine/Threonine Phosphatase PP2A. Immunity 13, 313–
738 322. doi:10.1016/S1074-7613(00)00031-5
739 Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M., 2005. Blast2GO:
740 a universal tool for annotation, visualization and analysis in functional genomics
741 research. Bioinforma. Appl. Note 21, 3674–3676. doi:10.1093/bioinformatics/bti610

742 Costa, M.M., Saraceni, P.R., Forn-Cuní, G., Dios, S., Romero, A., Figueras, A., Novoa, B.,
743 2013. IL-22 is a key player in the regulation of inflammation in fish and involves innate
744 immune cells and PI3K signaling. Dev. Comp. Immunol. 41, 746–755.
745 doi:10.1016/j.dci.2013.08.021
746 Creelman, E.K., Hauser, L., Simmons, R.K., Templin, W.D., Seeb, L.W., 2011. Temporal
747 and Geographic Genetic Divergence: Characterizing Sockeye Salmon Populations in the
748 Chignik Watershed, Alaska, Using Single-Nucleotide Polymorphisms. Trans. Am. Fish.
749 Soc. 140, 749–762. doi:10.1080/00028487.2011.584494
750 Dionne, M., Miller, K.M., Dodson, J.J., Bernatchez, L., 2009. MHC standing genetic
751 variation and pathogen resistance in wild Atlantic salmon. Philos. Trans. R. Soc. B Biol.
752 Sci. 364, 1555–1565. doi:10.1098/rstb.2009.0011
753 Dionne, M., Miller, K.M., Dodson, J.J., Caron, F., Bernatchez, L., 2007. Clinal variation in
754 MHC diversity with temperature: evidence for the role of host-pathogen interaction on
755 local adaptation in Atlantic salmon. Evolution 61, 2154–64. doi:10.1111/j.1558-
756 5646.2007.00178.x
757 Elphinstone, M.S., Hinten, G.N., Anderson, M.J., Nock, C.J., 2003. An inexpensive and high-
758 throughput procedure to extract and purify total genomic DNA for population studies.
759 Mol. Ecol. Notes 3, 317–320. doi:10.1046/j.1471-8286.2003.00397.x
760 Faure, M., Lafont, F., 2013. Pathogen-induced autophagy signaling in innate immunity. J.
761 Innate Immun. 5, 456–470. doi:10.1159/000350918
762 Flatman, P.W., 2002. Regulation of Na-K-2Cl cotransport by phosphorylation and protein-
763 protein interactions. Biochim. Biophys. Acta 1566, 140–151. doi:10.1016/S0005-
764 2736(02)00586-2
765 Foll, M., Gaggiotti, O., 2008. A genome-scan method to identify selected loci appropriate for
766 both dominant and codominant markers: a Bayesian perspective. Genetics 180, 977–
767 993. doi:10.1534/genetics.108.092221
768 Freeman, S.A., Grinstein, S., 2014. Phagocytosis: receptors, signal integration, and the
769 cytoskeleton. Immunol. Rev. 262, 193–215.
770 Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L., Nielsen, R., 2011.
771 Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main
772 Selective Pressure through Human Evolution. PLoS Genet. 7, 1–14.
773 doi:10.1371/journal.pgen.1002355
774 Funder, S., Demidov, I., Yelovicheva, Y., 2002. Hydrography and mollusc faunas of the
775 Baltic and the White Sea-North Sea seaway in the Eemian. Palaeogeogr. Palaeoclimatol.
776 Palaeoecol. 184, 275–304.
777 Gilbey, J., Verspoor, E., Mo, T.A., Sterud, E., Olstad, K., Hytterød, S., Jones, C., Noble, L.,
778 2006. Identification of genetic markers associated with Gyrodactylus salaris resistance
779 in Atlantic salmon Salmo salar. Dis. Aquat. Organ. 71, 119–129.
780 doi:10.3354/dao071119
781 Gillooly, J.F., Brown, J.H., West, G.B., Savage, V.M., Charnov, E.L., 2001. Effects of Size
782 and Temperature on Metabolic Rate. Science (80-.). 293, 2248–2251.
783 Goetzl, E.J., Wang, W., McGiffert, C., Huang, M.C., Gräler, M.H., 2004. Sphingosine 1-
784 phosphate and its G protein-coupled receptors constitute a multifunctional
785 immunoregulatory system. J. Cell. Biochem. 92, 1104–1114. doi:10.1002/jcb.20053
786 Günther, T., Coop, G., 2013. Robust identification of local adaptation from allele frequencies.
787 Genetics 195, 205–220. doi:10.1534/genetics.113.152462
788 Guo, B., Li, Z., Merilä, J., 2016. Population genomic evidence for adaptive differentiation in
789 the Baltic Sea herring. Mol. Ecol. 25, 2833–2852. doi:10.1111/mec.13657
790 Haarder, S., Kania, P.W., Bahlool, Q.Z.M., Buchmann, K., 2013. Expression of immune
791 relevant genes in rainbow trout following exposure to live Anisakis simplex larvae. Exp.
792 Parasitol. 135, 564–569. doi:10.1016/j.exppara.2013.09.011

Haasl, R.J., Payseur, B.A., 2016. Fifteen years of genomewide scans for selection: trends , lessons and unaddressed genetic sources of complication. Mol. Ecol. 25, 5–23. doi:10.1111/mec.13339

Hanington, P.C., Tam, J., Katzenback, B.A., Hitchen, S.J., Barreda, D.R., Belosevic, M., 2009. Development of macrophages of cyprinid fish. Dev. Comp. Immunol. 33, 411–429. doi:10.1016/j.dci.2008.11.004

Harbige, L.S., 2003. Fatty acids, the immune response, and autoimmunity: a question of n-6 essentiality and the balance between n-6 and n-3. Lipids 38, 323–341.

Hedrick, M.S., Hillman, S.S., Drewes, R.C., Withers, P.C., 2013. Lymphatic regulation in nonmammalian vertebrates. J. Appl. Physiol. 115, 297–308. doi:10.1152/japplphysiol.00201.2013

Heinz, L.X., Baumann, C.L., Köberlin, M.S., Snijder, B., Gawish, R., Shui, G., Sharif, O., Aspalter, I.M., Müller, A.C., Kandasamy, R.K., Breitwieser, F.P., Pichlmair, A., Bruckner, M., Rebsamen, M., Blüml, S., Karonitsch, T., Fauster, A., Colinge, J., Bennett, K.L., Knapp, S., Wenk, M.R., Superti-Furga, G., 2015. The Lipid-Modifying Enzyme SMPDL3B Negatively Regulates Innate Immunity. Cell Rep. 11, 1919–1928. doi:10.1016/j.celrep.2015.05.006

Hiroi, J., Yasumasu, S., Mccormick, S.D., Hwang, P.-P., Kaneko, T., 2008. Evidence for an apical Na–Cl cotransporter involved in ion uptake in a teleost fish. J. Exp. Biol. 211, 2584–2599. doi:10.1242/jeb.018663

Hoban, S., Kelley, J.L., Lotterhos, K.E., Antolin, M.F., Bradburd, G., Lowry, D.B., Poss, M.L., Reed, L.K., Storfer, A., Whitlock, M.C., 2016. Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. Am. Nat. 188, 379–397. doi:10.1086/688018

Hoffmann, E.K., Hoffmann, E., Lang, F., Zadunaisky, J.A., 2002. Control of Cl- transport in the operculum epithelium of Fundulus heteroclitus: long- and short-term salinity adaptation. Biochim. Biophys. Acta 1566, 129–139.

Holm, H.J., Skugor, S., Bjelland, A.K., Radunovic, S., Wadsworth, S., Koppang, E.O., Evensen, Ø., 2017. Contrasting expression of immune genes in scaled and scaleless skin of Atlantic salmon infected with young stages of Lepeophtheirus salmonis. Dev. Comp. Immunol. 67, 153–165. doi:10.1016/j.dci.2016.10.008

Jensen, L.F., Hansen, M.M., Carlsson, J., Loeschcke, V., Mensberg, K.L.D., 2005. Spatial and temporal genetic differentiation and effective population size of brown trout (Salmo trutta, L.) in small Danish rivers. Conserv. Genet. 6, 615–621. doi:10.1007/s10592-005-9014-8

Johnsen, B.O., Jensen, A.J., 1991. The Gyrodactylus story in Norway. Aquaculture 98, 289–302.

Kadiri, S., Monnier, C., Ganbold, M., Ledent, T., Capeau, J., Antoine, B., 2015. The nuclear retinoid-related orphan receptor-alpha regulates adipose tissue glyceroneogenesis in addition to hepatic gluconeogenesis. Am. J. Physiol. Endocrinol. Metab. 309, 105–114. doi:10.1152/ajpendo.00518.2014

Kaltz, O., Shykoff, J. a, 1998. Local adaptation in host-parasite systems. Heredity (Edinb). 81, 361–370. doi:10.1046/j.1365-2540.1998.00435.x

Kang, C., Tsai, H., Liu, C., Lee, T., Hwang, P.-P., 2010. Salinity-dependent expression of a Na+, K+, 2Cl− cotransporter in gills of the brackish medaka Oryzias dancena: A molecular correlate for hyposmoregulatory endurance. Comp. Biochem. Physiol. Part A 157, 7–18. doi:10.1016/j.cbpa.2010.05.013

Kania, P.W., Evensen, O., Larsen, T.B., Buchmann, K., 2010. Molecular and immunohistochemical studies on epidermal responses in Atlantic salmon Salmo salar L. induced by Gyrodactylus salaris Malmberg, 1957. J. Helminthol. 84, 166–172. doi:10.1017/S0022149X09990460

Kiiskinen, P., Huuskonen, H., Hyvärinen, H., Piironen, J., 2003. Effects of daylength and winter fasting on growth and smolting of one-year-old Saimaa landlocked salmon (Salmo salar m. sebago Girard) under fish farm conditions. Ann. Zool. Fennici 40, 441–458.

Kjærner-Semb, E., Ayllon, F., Furmanek, T., Wennevik, V., Dahle, G., Niemelä, E., Ozerov, M., Vähä, J.-P., Glover, K.A., Rubin, C.J., Wargelius, A., Edvardsen, R.B., 2016. Atlantic salmon populations reveal adaptive divergence of immune related genes - a duplicated genome under selection. BMC Genomics 17, 1–12. doi:10.1186/s12864-016-2867-z

Krasnov, A., Skugor, S., Todorcevic, M., Glover, K.A., Nilsen, F., 2012. Gene expression in Atlantic salmon skin in response to infection with the parasitic copepod Lepeophtheirus salmonis, cortisol implant, and their combination. BMC Genomics 13, 1–15. doi:10.1186/1471-2164-13-130

Krasnov, A., Timmerhaus, G., Schiøtz, B.L., Torgersen, J., Afanasyev, S., Iliev, D., Jørgensen, J., Takle, H., Jørgensen, S.M., 2011. Genomic survey of early responses to viruses in Atlantic salmon, Salmo salar L. Mol. Immunol. 49, 163–174. doi:10.1016/j.molimm.2011.08.007

Kudersky, L., Ieshko, E., Schulman, B., 2003. Distribution and range formation history of the monogenean Gyrodactylus salaris Malmberg, 1957 — a parasite of juvenile Atlantic salmon Salmo salar Linnaeus, in: Atlantic Salmon Biology, Conservation and Restoration. pp. 77–83.

Kumari, J., Larsen, A.N., Bogwald, J., Dalmo, R.A., 2009. Interleukin-17D in Atlantic salmon (Salmo salar): molecular characterization, 3D modelling and promoter analysis. Fish Shellfish Immunol. 27, 647–659. doi:10.1016/j.fsi.2009.08.009

Kusakabe, M., Ishikawa, A., Ravinet, M., Yoshida, K., Makino, T., Toyoda, A., Fujiyama, A., Kitano, J., 2017. Genetic basis for variation in salinity tolerance between stickleback ecotypes. Mol. Ecol. 26, 304–319. doi:10.1111/mec.13875

Kuusela, J., Holopainen, R., Meinilä, M., Anttila, P., Koski, P., Ziętara, M.S., Veselov, A., Primmer, C.R., Lumme, J., 2009. Clonal structure of salmon parasite Gyrodactylus salaris on a coevolutionary gradient on Fennoscandian salmon (Salmo salar). Ann. Zool. Fennici 46, 21–33.

Kuusela, J., Zietara, M.S., Lumme, J., 2007. Hybrid origin of Baltic salmon-specific parasite Gyrodactylus salaris: a model for speciation by host switch for hemiclonal organisms. Mol. Ecol. 16, 5234–45. doi:10.1111/j.1365-294X.2007.03562.x

Li, G., Zhao, Y., Wang, J., Liu, B., Sun, X., Guo, S., Feng, J., 2017. Transcriptome profiling of developing spleen tissue and discovery of immune-related genes in grass carp (Ctenopharyngodon idella). Fish Shellfish Immunol. 60, 400–410. doi:10.1016/j.fsi.2016.12.012

Li, J., Li, Y., Cao, Y., Yuan, M., Gao, Z., Guo, X., Zhu, F., Wang, Y., Xu, J., 2014. Polycomb chromobox (Cbx) 7 modulates activation-induced CD4+ T cell apoptosis. Arch. Biochem. Biophys. 564, 184–188. doi:10.1016/j.abb.2014.10.004

Lien, S., Koop, B.F., Sandve, S.R., Miller, J.R., Kent, M.P., Nome, T., Hvidsten, T.R., Leong, J.S., Minkley, D.R., Zimin, A., Grammes, F., Grove, H., Gjuvsland, A., Walenz, B., Hermansen, R.A., von Schalburg, K., Rondeau, E.B., Di Genova, A., Samy, J.K.A., Olav Vik, J., Vigeland, M.D., Caler, L., Grimholt, U., Jentoft, S., Inge Våge, D., de Jong, P., Moen, T., Baranski, M., Palti, Y., Smith, D.R., Yorke, J.A., Nederbragt, A.J., Tooming-Klunderud, A., Jakobsen, K.S., Jiang, X., Fan, D., Hu, Y., Liberles, D.A., Vidal, R., Iturra, P., Jones, S.J.M., Jonassen, I., Maass, A., Omholt, S.W., Davidson, W.S., 2016. The Atlantic salmon genome provides insights into rediploidization. Nature 533, 200–205. doi:10.1038/nature17164

Lillico, D.M.E., Zwozdesky, M.A., Pemberton, J.G., Deutscher, J.M., Jones, L.O., Chang,

J.P., Stafford, J.L., 2017. Teleost leukocyte immune-type receptors activate distinct phagocytic modes for target acquisition and engulfment. J. Leukoc. Biol. 98, 235–248. doi:10.1189/jlb.2A0215-039RR
Lim, J., Wiedemann, A., Tzircotis, G., Monkley, S.J., Critchley, D.R., Caron, E., 2007. An Essential Role for Talin during aMB2-mediated Phagocytosis. Mol. Biol. Cell 18, 976–985. doi:10.1091/mbc.E06
Limborg, M.T., Blankenship, S.M., Young, S.F., Utter, F.M., Seeb, L.W., Hansen, M.H.H., Seeb, J.E., 2012a. Signatures of natural selection among lineages and habitats in Oncorhynchus mykiss. Ecol. Evol. 2, 1–18. doi:10.1002/ece3.59
Limborg, M.T., Helyar, S.J., De Bruyn, M., Taylor, M.I., Nielsen, E.E., Ogden, R., Carvalho, G.R., Bekkevold, D., 2012b. Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (Clupea harengus). Mol. Ecol. 21, 3686–703. doi:10.1111/j.1365-294X.2012.05639.x
Lindenstrom, T., Sigh, J., Dalgaard, M.B., Buchmann, K., 2006. Skin expression of IL-1b in East Atlantic salmon, Salmo salar L., highly susceptible to Gyrodactylus salaris infection is enhanced compared to a low susceptibility Baltic stock. J. Fish Dis. 29, 123–128.
Lionetto, M.G., Schettino, T., 2006. The Na+ K+ 2Cl- cotransporter and the osmotic stress response in a model salt transport epithelium. Acta Physiol. 187, 115–124. doi:10.1111/j.1748-1716.2006.01536.x
Lumme, J., Ozerov, M.Y., Veselov, A.E., Primmer, C.R., 2016. The Formation of Landlocked Populations of Atlantic Salmon, in: Evolutionary Biology of the Atlantic Salmon. pp. 26–43.
Mace, E.M., Monkley, S.J., Critchley, D.R., Takei, F., 2009. A Dual Role for Talin in NK Cell Cytotoxicity: Activation of LFA-1-Mediated Cell Adhesion and Polarization of NK Cells. J. Immunol. 182, 948–956. doi:10.4049/jimmunol.182.2.948
Makrinos, D.L., Bowden, T.J., 2016. Natural environmental impacts on teleost immune function. Fish Shellfish Immunol. 53, 50–57. doi:10.1016/j.fsi.2016.03.008
Matejusová, I., Felix, B., Sorsa-Leslie, T., Gilbey, J., Noble, L.R., Jones, C.S., Cunningham, C.O., 2006. Gene expression profiles of some immune relevant genes from skin of susceptible and responding Atlantic salmon (Salmo salar L.) infected with Gyrodactylus salaris (Monogenea) revealed by suppressive subtractive hybridisation. Int. J. Parasitol. 36, 1175–1183. doi:10.1016/j.ijpara.2006.04.009
Medzhitov, R., 2007. Recognition of microorganisms and activation of the immune response. Nature 449, 819–826. doi:10.1038/nature06246
Mikrou, A., Marioli, D., Papanastasiou, A.D., Zarkadis, I.K., 2009. CR3 complement receptor: Cloning and characterization in rainbow trout. Fish Shellfish Immunol. 26, 19–28. doi:10.1016/j.fsi.2008.10.005
Mitchell, P.S., Emerman, M., Malik, H.S., 2013. An evolutionary perspective on the broad antiviral specificity of MxA. Curr. Opin. Microbiol. 16, 493–499. doi:10.1016/j.mib.2013.04.005
Moen, T., Baranski, M., Sonesson, A.K., Kjøglum, S., 2009. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (Salmo salar): population-level associations between markers and trait. BMC Genomics 10, 1–14. doi:10.1186/1471-2164-10-368
Moen, T., Torgersen, J., Santi, N., Davidson, W.S., Baranski, M., Ødegård, J., Kjøglum, S., Velle, B., Kent, M., Lubieniecki, K.P., Isdal, E., Lien, S., 2015. Epithelial Cadherin Determines Resistance to Infectious Pancreatic Necrosis Virus in Atlantic Salmon. Genetics 200, 1313–1326. doi:10.1534/genetics.115.175406
Nakamura, T., Liu, Y., Hiratal, D., Nambal, H., Harada, S., Hirokawa, T., Miyakawa, T., 1993. Protein phosphatase type 2B (calcineurin)-mediated, FK506-sensitive regulation

of intracellular ions in yeast is an important determinant for adaptation to high salt stress conditions. EMBO J. 12, 4063–4071.
Nakanishi, T., Shibasaki, Y., Matsuura, Y., 2015. T Cells in Fish. Biology (Basel). 4, 640–663. doi:10.3390/biology4040640
Narum, S.R., Zendt, J.S., Frederiksen, C., Campbell, N., Matala, A., Sharp, W.R., 2011. Candidate genetic markers associated with anadromy in Oncorhynchus mykiss of the Klickitat River. Trans. Am. Fish. Soc. 140, 843–854. doi:10.1080/00028487.2011.588131
Naumenko, M.A., Karetnikov, S.G., Tikhomirov, A.I., 1996. Main features of the thermal regime of Lake Ladoga during the ice-free period. Hydrobiologia 322.
Nielsen, E.E., Hemmer-Hansen, J., Poulsen, N.A., Loeschcke, V., Moen, T., Johansen, T., Mittelholzer, C., Taranger, G., Ogden, R., Carvalho, G.R., 2009. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (Gadus morhua). BMC Evol. Biol. 9, 1–11. doi:10.1186/1471-2148-9-276
Nielsen, R., Williamson, S., Kim, Y., Hubisz, M.J., Clark, A.G., Bustamante, C., 2005. Genomic scans for selective sweeps using SNP data. Genome Res. 15, 1566–75. doi:10.1101/gr.4252305
Nilsen, T.O., Ebbesson, L.O.E., Kiilerich, P., Björnsson, B.T., Madsen, S.S., McCormick, S.D., Stefansson, S.O., 2008. Endocrine systems in juvenile anadromous and landlocked Atlantic salmon (Salmo salar): seasonal development and seawater acclimation. Gen. Comp. Endocrinol. 155, 762–72. doi:10.1016/j.ygcen.2007.08.006
Nilsen, T.O., Ebbesson, L.O.E., Stefansson, S.O., 2003. Smolting in anadromous and landlocked strains of Atlantic salmon (Salmo salar). Aquaculture 222, 71–82. doi:10.1016/S0044-8486(03)00103-0
Nilsson, J., Gross, R., Asplund, T., Dove, O., Jansson, H., Kelloniemi, J., Kohlmann, K., Löytynoja, A., Nielsen, E.E., Paaver, T., Primmer, C.R., Titov, S., Vasemägi, A., Veselov, A., Ost, T., Lumme, J., 2001. Matrilinear phylogeography of Atlantic salmon (Salmo salar L.) in Europe and postglacial colonization of the Baltic Sea area. Mol. Ecol. 10, 89–102.
Oleksyk, T.K., Smith, M.W., O'Brien, S.J., 2010. Genome-wide scans for footprints of natural selection. Philos. Trans. R. Soc. B Biol. Sci. 365, 185–205. doi:10.1098/rstb.2009.0219
Oleksyk, T.K., Zhao, K., De La Vega, F.M., Gilbert, D.A., O'Brien, S.J., Smith, M.W., 2008. Identifying selected regions from heterozygosity and divergence using a light-coverage genomic dataset from two human populations. PLoS One 3, 1–15. doi:10.1371/journal.pone.0001712
Orr, H.A., 2000. Adaptation and the cost of complexity. Evolution (N. Y). 54, 13–20.
Oshiumi, H., Sakai, K., Matsumoto, M., Seya, T., 2010. DEAD/H BOX 3 (DDX3) helicase binds the RIG-I adaptor IPS-1 to up-regulate IFN-B-inducing potential. Eur. J. Immunol. 40, 940–948. doi:10.1002/eji.200940203
Ozerov, M., Veselov, A.E., Lumme, J., Primmer, C.R., 2013. Temporal variation of genetic composition in Atlantic salmon populations from the Western White Sea Basin: influence of anthropogenic factors? BMC Genet. 14, 1–15. doi:10.1186/1471-2156-14-88
Ozerov, M.Y., Veselov, A.E., Lumme, J., Primmer, C.R., 2012. "Riverscape" genetics: river characteristics influence the genetic structure and diversity of anadromous and freshwater Atlantic salmon (Salmo salar) populations in northwest Russia. Can. J. Fish. Aquat. Sci. 69, 1–12. doi:10.1139/f2012-114
Ozerov, M.Y., Veselov, A.J., Lumme, J., Primmer, C.R., 2010. Genetic structure of freshwater Atlantic salmon (Salmo salar L.) populations from the lakes Onega and Ladoga of northwest Russia and implications for conservation. Conserv. Genet. 11,

997 1711–1724. doi:10.1007/s10592-010-0064-1
998 Palstra, F.P., Ruzzante, D.E., 2010. A temporal perspective on population structure and gene
999 flow in Atlantic salmon (Salmo salar) in Newfoundland, Canada. Can. J. Fish. Aquat.
1000 Sci. 67, 225–242. doi:10.1139/F09-176
1001 Pankratz, V.S., Vierkant, R.A., Byrne, M.M.O., Ovsyannikova, I.G., Poland, G.A., 2010.
1002 Associations between SNPs in candidate immune-relevant genes and rubella antibody
1003 levels: a multigenic assessment. BMC Immunol. 11, 1–10.
1004 Papakostas, S., Vøllestad, L.A., Bruneaux, M., Aykanat, T., Vanoverbeke, J., Ning, M.,
1005 Primmer, C.R., Leder, E.H., 2014. Gene pleiotropy constrains gene expression changes
1006 in fish adapted to different thermal conditions. Nat. Commun. 5, 1–9.
1007 doi:10.1038/ncomms5071
1008 Park, B., Buti, L., Lee, S., Matsuwaki, T., Spooner, E., Brinkmann, M.M., Nishihara, M.,
1009 Ploegh, H.L., 2011. Granulin Is a Soluble Cofactor for Toll-like Receptor 9 Signaling.
1010 Immunity 34, 505–513. doi:10.1016/j.immuni.2011.01.018
1011 Peng, J., Larondelle, Y., Pham, D., Ackman, R.G., Rollin, X., 2003. Polyunsaturated fatty
1012 acid profiles of whole body phospholipids and triacylglycerols in anadromous and
1013 landlocked Atlantic salmon (Salmo salar L .) fry. Comp. Biochem. Physiol. 134, 335–
1014 348. doi:10.1016/S1096-4959(02)00263-4
1015 Pritchard, V.L., Erkinaro, J., Kent, M.P., Niemelä, E., Orell, P., Lien, S., Primmer, C.R.,
1016 2016. Single nucleotide polymorphisms to discriminate different classes of hybrid
1017 between wild Atlantic salmon and aquaculture escapees. Evol. Appl. 9, 1017–1031.
1018 doi:10.1111/eva.12407
1019 Przeworski, M., Coop, G., Wall, J.D., 2005. The signature of positive selectionon standing
1020 genetic variation. Evolution (N. Y). 59, 2312–2323.
1021 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J.,
1022 Sklar, P., de Bakker, P.I.W., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-
1023 genome association and population-based linkage analyses. Am. J. Hum. Genet. 81,
1024 559–575. doi:10.1086/519795
1025 Quinlan, A.R., Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing
1026 genomic features. Bioinformatics 26, 841–842. doi:10.1093/bioinformatics/btq033
1027 Råberg, L., Sim, D., Read, A.F., 2007. Disentangling genetic variation for resistance and
1028 tolerance to infectious diseases in animals. Science (80-.). 318, 812–814.
1029 doi:10.1126/science.1148526
1030 R Core Team, 2016. A language and environment for statistical computing. R Foundation for
1031 Statistical Computing, Vienna, Austria. https://www.r-project.org/.
1032 Rebl, A., Goldammer, T., Seyfert, H.-M., 2010. Toll-like receptor signaling in bony fish. Vet.
1033 Immunol. Immunopathol. 134, 139–150. doi:10.1016/j.vetimm.2009.09.021
1034 Reyes-Lopez, F.E., Romeo, J.S., Vallejos-Vidal, E., Reyes-Cerpa, S., Sandino, A.M., Tort,
1035 L., Mackenzie, S., Imarai, M., 2015. Differential immune gene expression profiles in
1036 susceptible and resistant full-sibling families of Atlantic salmon (Salmo salar)
1037 challenged with infectious pancreatic necrosis virus (IPNV). Dev. Comp. Immunol. 53,
1038 210–221. doi:10.1016/j.dci.2015.06.017
1039 Saarnisto, M., Saarinen, T., 2001. Deglaciation chronology of the Scandinavian Ice Sheet
1040 from the Lake Onega Basin to the Salpausselkä End Moraines. Glob. Planet. Change 31,
1041 387–405.
1042 Sadeghi, H., Gupta, Y., Möller, S., Samavedam, U.K., Behnen, M., Kasprick, A., Bieber, K.,
1043 Müller, S., Kalies, K., de Castro Marques, A., Recke, A., Schmidt, E., Zillikens, D.,
1044 Laskay, T., Mariani, J., Ibrahim, S.M., Ludwig, R.J., 2015. The retinoid-related orphan
1045 receptor alpha is essential for the end-stage effector phase of experimental epidermolysis
1046 bullosa acquisita. J. Pathol. 237, 111–122. doi:10.1002/path.4556
1047 Sancho, D., Vicente-Manzanares, M., Mittelbrunn, M., Montoya, M.C., Gordon-Alonso, M.,

Serrador, J.M., Sanchez-Madrid, F., 2002. Regulation of microtubule- organizing center orientation and actomyosin cytoskeleton rearrangement during immune interactions. Immunol. Rev. 189, 84–97.
Schiöth, H.B., Fredriksson, R., 2005. The GRAFS classification system of G-protein coupled receptors in comparative perspective. Gen. Comp. Endocrinol. 142, 94–101. doi:10.1016/j.ygcen.2004.12.018
Schlötterer, C., 2003. Hitchhiking mapping - functional genomics from the population genetics perspective. Trends Genet. 19, 32–38.
Schröder, M., 2011. Viruses and the human DEAD-box helicase DDX3: inhibition or exploitation? Biochem. Soc. Trans. 39, 679–683. doi:10.1042/BST0390679
Shiozaki, K., Russell, P., 1995. Counteractive roles of protein phosphatase 2C (PP2C) and a MAP kinase kinase homolog in the osmoregulation of fission yeast. EMBO J. 14, 492–502.
Smit, M.J., Vischer, H.F., Bakker, R.A., Jongejan, A., Timmerman, H., Pardo, L., Leurs, R., 2007. Pharmacogenomic and structural analysis of constitutive G protein-coupled receptor activity. Annu. Rev. Pharmacol. Toxicol. 47, 53–87. doi:10.1146/annurev.pharmtox.47.120505.105126
Smout, M.J., Laha, T., Mulvenna, J., Sripa, B., Suttiprapa, S., Jones, A., Brindley, P.J., Loukas, A., 2009. A granulin-like growth factor secreted by the carcinogenic liver fluke, Opisthorchis viverrini, promotes proliferation of host cells. PLoS Pathog. 5, 1–16. doi:10.1371/journal.ppat.1000611
Sommer, S., 2005. The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front. Zool. 2, 1–18. doi:10.1186/1742-9994-2-16
Stanton, R.J., Prod'homme, V., Purbhoo, M.A., Moore, M., Aicheler, R.J., Heinzmann, M., Bailer, S.M., Haas, J., Antrobus, R., Weekes, M.P., Lehner, P.J., Vojtesek, B., Miners, K.L., Man, S., Wilkie, G.S., Davison, A.J., Wang, E.C.Y., Tomasec, P., Wilkinson, G.W.G., 2014. HCMV pUL135 Remodels the Actin Cytoskeleton to Impair Immune Recognition of Infected Cells. Cell Host Microbe 16, 201–214. doi:10.1016/j.chom.2014.07.005
Sun, Y., Liu, C., Paul, J., Evans, L.P., Tian, K.T., Zhang, B., Stahl, A., Pu, W.T., Kamenecka, T.M., Solt, L.A., Chen, J., 2015. Nuclear receptor RORα regulates pathologic retinal angiogenesis by modulating SOCS3-dependent inflammation. PNAS 112, 10401–10406. doi:10.1073/pnas.1504387112
Sunyer, J.O., 2013. Fishing for mammalian paradigms in the teleost immune system. Nat. Immunol. 14, 1–19. doi:10.1007/s11103-011-9767-z.Plastid
Tadiso, T.M., Krasnov, A., Skugor, S., Afanasyev, S., Hordvik, I., Nilsen, F., 2011. Gene expression analyses of immune responses in Atlantic salmon during early stages of infection by salmon louse (Lepeophtheirus salmonis) revealed bi-phasic responses coinciding with the copepod-chalimus transition. BMC Genomics 12, 1–17.
Tolstikov, A. V., Petrov, M.P., 2006. Mean multiannual temperature regime of the surface water layer in the near-shore regions of the White Sea. Oceanology 46, 318–324. doi:10.1134/S0001437006030039
Tonteri, A., Titov, S., Veselov, A., Zubchenko, A., Koskinen, M.T., Lesbarrères, D., Kaluzchin, S., Bakhmet, I., Lumme, J., Primmer, C.R., 2005. Phylogeography of anadromous and non-anadromous Atlantic salmon (Salmo salar) from northern Europe. Ann. Zool. Fennici 41, 1–22.
Tonteri, A., Vasemägi, A., Lumme, J., Primmer, C.R., 2010. Beyond MHC: signals of elevated selection pressure on Atlantic salmon (Salmo salar) immune-relevant loci. Mol. Ecol. 19, 1273–1282. doi:10.1111/j.1365-294X.2010.04573.x
Tonteri, A., Vasemägi, A., Lumme, J., Primmer, C.R., 2008. Use of differential expression data for identification of novel immune relevant expressed sequence tag-linked

microsatellite markers in Atlantic salmon (Salmo salar L.). Mol. Ecol. Resour. 8, 1486–1490. doi:10.1111/j.1755-0998.2008.02317.x

Tonteri, A., Veselov, A.J., Titov, S., Lumme, J., Primmer, C.R., 2007. The effect of migratory behaviour on genetic diversity and population divergence: a comparison of anadromous and freshwater Atlantic salmon Salmo salar. J. Fish Biol. 70, 381–398. doi:10.1111/j.1095-8649.2007.01519.x

Tonteri, A., Veselov, A.J., Zubchenko, A. V, Lumme, J., Primmer, C.R., 2009. Microsatellites reveal clear genetic boundaries among Atlantic salmon (Salmo salar) populations from the Barents and White seas, northwest Russia. Can. J. Fish. Aquat. Sci. 66, 717–735. doi:10.1139/F09-010

Trombetta, E.S., Ebersold, M., Garrett, M.P., Mellman, I., 2003. Activation of Lysosomal Function During Dendritic Cell Maturation. Science (80-.). 299, 1400–1403. doi:10.1126/science.1080106

Vähä, J.-P., Erkinaro, J., Niemelä, E., Primmer, C.R., 2008. Temporally stable genetic structure and low migration in an Atlantic salmon population complex: implications for conservation and management. Evol. Appl. 1, 137–154. doi:10.1111/j.1752-4571.2007.00007.x

Vasemägi, A., Nilsson, J., Primmer, C.R., 2005. Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in atlantic salmon (Salmo salar L.). Mol. Biol. Evol. 22, 1067–76. doi:10.1093/molbev/msi093

Vilas, R., Vandamme, S.G., Vera, M., Bouza, C., Maes, G.E., Volckaert, F.A.M., Martínez, P., 2015. A genome scan for candidate genes involved in the adaptation of turbot (Scophthalmus maximus). Mar. Genomics 23, 77–86. doi:10.1016/j.margen.2015.04.011

Wagner, G.P., Zhang, J., 2011. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nat. Rev. Genet. 12, 204–213. doi:10.1038/nrg2949

Wang, T., Diaz-Rosales, P., Costa, M.M., Campbell, S., Snow, M., Collet, B., Martin, S.A.M., Secombes, C.J., 2011. Functional Characterization of a Nonmammalian IL-21: Rainbow Trout Oncorhynchus mykiss IL-21 Upregulates the Expression of the Th Cell Signature Cytokines IFN-γ, IL-10, and IL-22. J. Immunol. 186, 708–721. doi:10.4049/jimmunol.1001203

Watkins, G.R., Wang, N., Mazalouskas, M.D., Gomez, R.J., Guthrie, C.R., Kraemer, B.C., Schweiger, S., Spiller, B.W., Wadzinski, B.E., 2012. Monoubiquitination promotes calpain cleavage of the protein phosphatase 2A (PP2A) regulatory subunit a4, altering PP2A stability and microtubule-associated protein phosphorylation. J. Biol. Chem. 287, 24207–24215. doi:10.1074/jbc.M112.368613

Whitlock, M.C., Lotterhos, K.E., 2015. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of Fst*. Am. Nat. 186, 26–36. doi:10.1086/682949

Winder, M., Schindler, D.E., 2004. Climate change uncouples trophic interactions in an Aquatic ecosystem. Ecology 85, 2100–2106.

Yamamoto, H., Hatano, M., Iitsuka, Y., Mahyar, N.-S., Yamamoto, M., Tokuhisa, T., 1995. Two forms of Hox11, a T cell leukemia oncogene, are expressed in fetal spleen but not in primary lymphocytes. Mol. Immunol. 32, 1177–1182. doi:10.1016/0161-5890(95)00100-X

Yang, X., Downes, M., Yu, R.T., Bookout, A.L., He, W., Straume, M., Mangelsdorf, D.J., Evans, R.M., 2006. Nuclear Receptor Expression Links the Circadian Clock to Metabolism. Cell 126, 801–810. doi:10.1016/j.cell.2006.06.050

Yang, X.O., Pappu, B.P., Nurieva, R., Akimzhanov, A., Kang, H.S., Chung, Y., Ma, L., Shah, B., Panopoulos, A.D., Schluns, K.S., Watowich, S.S., Tian, Q., Jetten, A.M., Dong, C., 2008. T Helper 17 Lineage Differentiation Is Programmed by Orphan Nuclear

Receptors RORa and RORy. Immunity 28, 29–39. doi:10.1016/j.immuni.2007.11.016
Zapata, A., Diez, B., Cejalvo, T., Gutiérrez-De Frías, C., Cortés, A., 2006. Ontogeny of the immune system of fish. Fish Shellfish Immunol. 20, 126–136. doi:10.1016/j.fsi.2004.09.005
Zhu, L., Nie, L., Zhu, G., Xiang, L., Shao, J., 2013. Advances in research of fish immune-relevant genes: A comparative overview of innate and adaptive immunity in teleosts. Dev. Comp. Immunol. 39, 39–62. doi:10.1016/j.dci.2012.04.001
Zueva, K.J., Lumme, J., Veselov, A.E., Kent, M.P., Lien, S., Primmer, C.R., 2014. Footprints of Directional Selection in Wild Atlantic Salmon Populations: Evidence for Parasite-Driven Evolution? PLoS One 9, 1–16. doi:10.5061/dryad.gm367

Figure 1. Sampling site locations. Populations that were also studied previously (Zueva et al. 2014) are presented in grey, while white circles indicate populations that were added for this study.

(Figure 1 is intended as a 2-column fitting image; colour)

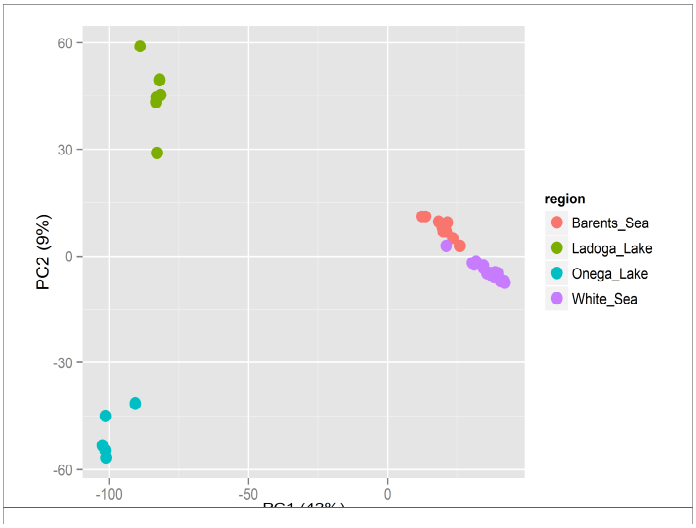

Figure 2. Population sub-structuring based on principal component (PC) analysis. Individual populations plotted as dots and coloured based on their geographic location; percentage of variance explained by PC1 and PC2 is given in brackets.
(Figure 2 is intended as a 1 or 1.5-column fitting image; colour)

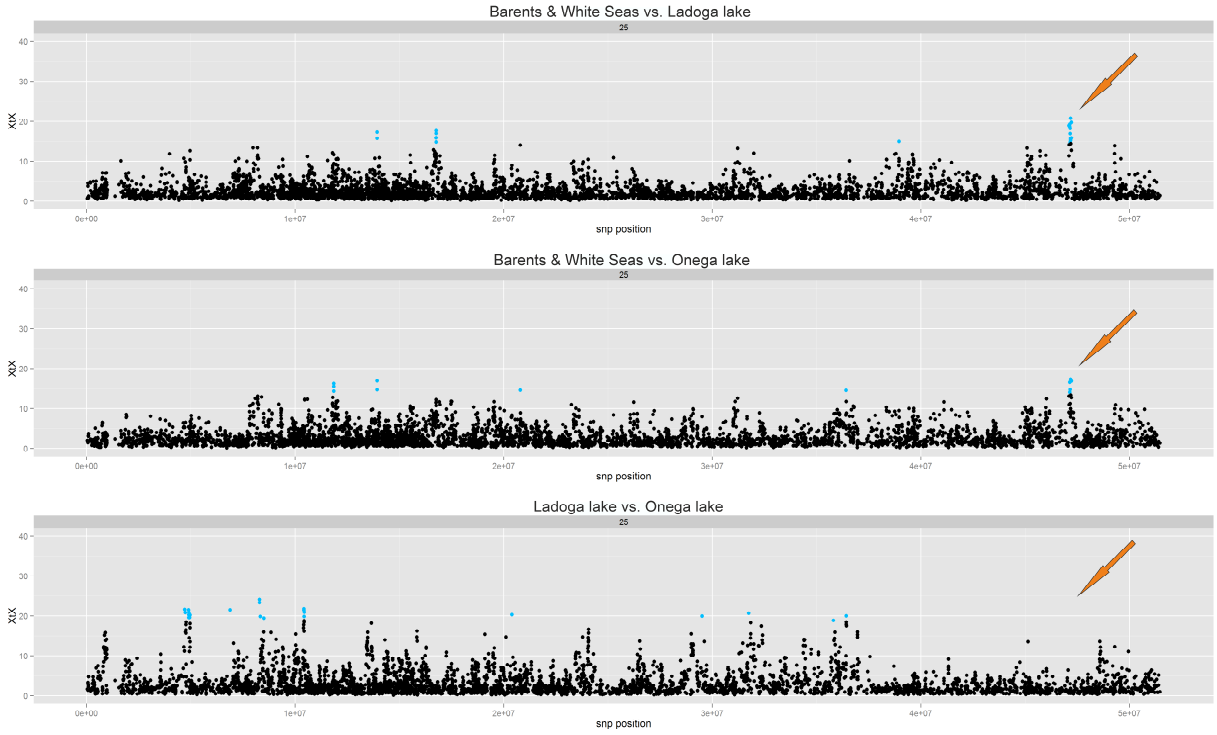

Figure 3. Levels of SNP differentiation (measured using the X^TX statistic, chromosome 25) between *G. salaris* susceptible (Barents & White Seas) and resistant (lakes Onega and Ladoga) salmon stocks. Each dot represents one SNP, and outlier SNPs with elevated X^TX (above 99.5% percentile) are marked with blue. Arrows indicate one candidate region associated with differences in parasite response, characterized by a high density of outlier SNPs in both "resistant vs susceptible" comparisons, but absent from the "resistant vs resistant" test.
(Figure 3 is intended as a 2-column fitting image; colour)

№	Population	Pool size	Sampling year	Coordinates	
Non-anadromous Ladoga lake					
1	Taipale	30	2000	60° 37’ 26.4’’ N,	30° 30’ 07.2’’ E
2	Hiitola	44	2006	61° 11’ 56.5’’ N,	29° 46’ 12.3’’ E
3	Sysky*	33	1999	61° 38’ 51.5’’ N,	31° 16’ 18.3’’ E
4	Uuksa	29	2006	61° 29’ 24.5’’ N,	31° 35’ 54.0’’ E
5	Tulema	63	2006	61° 21’ 25.0’’ N,	31° 50’ 28.4’’ E
6	Vidlitsa	44	2006	61° 10’ 32.7’’ N,	32° 23’ 12.1’’ E
Non-anadromous Onega lake					
7	Shuya	31	1996	61° 52’ 00’’ N,	34° 18’ 00’’ E
8	Lizma	23	1996	62° 22’ 35.6’’ N,	34° 30’ 18.9’’ E
9	Kumsa	31	2004	62° 54’ 31.4’’ N,	34° 28’ 17.5’’ E
10	Pyalma*	46	2001	62° 24’ 14.6’’ N,	35° 52’ 24.2’’ E
11	Tuba	40	2001	62° 15’ 00’’ N,	35° 49’ 18.9’’ E
Anadromous Barents Sea					
12	Tenojoki_1	326	2001-2003	69° 54’ 59.5’’ N,	27° 03’ 24.2’’ E
13	Tenojoki_2	137	2001-2003	69° 25’ 55.0’’ N,	25° 48’ 26.0’’ E
14	Näätämo	240	2006-2008	69° 42’ 27.9’’ N,	28° 59’ 16.6’’ E
15	Titovka	38	2000	69° 28’ 48.6’’ N,	31° 49’ 43.5’’ E
16	Z_Litsa	43	2000	69° 24’ 30.3’’ N,	32° 09’ 13.9E
17	Ura	44	2000	69° 17’ 29.7’’ N,	32° 49’ 27.0’’ E
18	Tuloma*	40	1998	68° 40’ 12.7’’ N,	31° 56’ 20.5’’ E
19	Kola	40	2000	68° 49’ 00’’ N,	33° 05’ 00’’ E
20	Drozdovka	48	2001	68° 17’ 29.1’’ N,	38° 26’ 27.2’’ E
21	Yokanga	39	2001	67° 59’ 54.4’’ N,	39° 42’ 38.4’’ E
Anadromous White Sea					
22	Kachovka	66	2008	67° 26’ 30.9’’ N,	40° 57’ 16.2’’ E
23	Ponoi	83	2008	67° 07’ 27.6’’ N,	40° 56’ 08.0’’ E
24	P_Lebyazia*	48	2001	67° 04’ 00’’ N,	38° 34’ 00’’ E
25	Danilovka	48	2008	66° 44’ 25.0’’ N,	41° 01’ 21.1’’ E
26	Sneznitsa	25	2008	66° 34’ 47.6’’ N,	40° 41’ 56.5’’ E
27	Sosnovka	47	2008	66° 30’ 33.2’’ N,	40° 35’ 19.7’’ E
28	Babya	25	2008	66° 23’ 16.0’’ N,	40° 17’ 25.2’’ E
29	Lihodeevka	53	2008	66° 20’ 09.1’’ N,	40° 10’ 46.5’’ E
30	Pulonga	57	2008	66° 15’ 58.3’’ N,	39° 58’ 18.4’’ E
31	Ust_Pyalka	45	2008	66° 12’ 00’’ N,	39° 30’ 00’’ E
32	Strelna	64	2008	66° 04’ 33.4’’ N,	38° 38’ 22.6’’ E
33	Chavanga	42	2008	66° 09’ 00’’ N,	37° 46’ 00’’ E
34	Yapoma*	34	2000	66° 37’ 25.2’’ N,	36° 12’ 10.0’’ E
35	Indera	60	2008	66° 14’ 30.7’’ N,	37° 08’ 43.2’’ E
36	Varzuga	48	2008	66° 24’ 00’’ N,	36° 37’ 00’’ E
37	Olenitsa	46	2000	66° 28’ 25.5’’ N,	35° 20’ 11.1’’ E
38	Umba	44	2001	66° 49’ 00’’ N,	34° 17’ 00’’ E
39	Nilma	39	2005	66° 30’ 04.3’’ N,	33° 08’ 04.3’’ E
40	Pongoma*	41	2005	65° 17’ 00’’ N,	34° 00’ 00’’ E
41	Suma*	36	1999	64° 16’ 58.9’’ N,	35° 24’ 08.5’’ E
42	SD_Emtsa*	42	2001	63° 30’ 36.9’’ N,	41° 50’ 19.6’’ E
43	Megra	36	2001	66° 09’ 24’’ N,	41° 34’ 44.1’’ E

Table 1. Details of the studied Atlantic salmon populations: regional grouping, location, and number of individuals pooled. Populations marked with asterisk (*) have been studied previously in Zueva et al. 2014.

Chromo-some	Gene name	Gene start, bp	Gene end, bp	Gene product
ssa01	LOC106569418	142150894	142165267	uncharacterized protein C18orf25-like
ssa01	LOC106608943	74119498	74124819	chromobox protein homolog 7-like
ssa01	LOC106609330	74197728	74204071	chondroadherin-like protein
ssa01	LOC106612532	93113743	93675639	CUB and sushi domain-containing protein 1-like
ssa03	LOC106599503"	19230012	19476371	partitioning defective 3 homolog
ssa04	LOC106602322	7837836	7849338	probable E3 ubiquitin-protein ligase HERC3
ssa04	pdk3	35428078	35436339	pyruvate dehydrogenase kinase 2C isoenzyme 3
ssa05	LOC106604946	40861196	40914104	fibroblast growth factor receptor-like 1
ssa05	**LOC106605373**	58379225	58400842	N-acetylglucosamine-1-phosphodiester α -N-acetylglucosaminidase (NAGPA)-like
ssa06	LOC106608134*"	72162107	72202063	serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit γ
ssa07	LOC106608623	13548236	13565188	phospholipase A-2-activating protein-like
ssa07	LOC106608629	13151483	13287518	serine/threonine-protein phosphatase 2B catalytic subunit α isoform
ssa09	LOC106611113	28405568	28417294	intron-binding protein aquarius-like
ssa09	LOC106611166	30445432	30494810	sodium/potassium/calcium exchanger 4-like
ssa09	LOC106611262	37711638	37752583	nesprin-3-like
ssa10	adgrl2	2603630	2817943	adhesion G protein-coupled receptor L2
ssa10	**ddx11**	86640661	86652856	DEAD/H (Asp-Glu-Ala-Asp/His) box helicase 11
ssa10	fkbp4	86653186	86669653	FK506 binding protein 4 2C 59kDa
ssa10	hipk3	100391425	100436959	homeodomain interacting protein kinase 3
ssa10	LOC106560916	76184088	76220764	sphingomyelin phosphodiesterase 3-like
ssa10	LOC106561152	83386500	83544652	talin-2-like
ssa10	usb1	75766859	75773108	U6 snRNA biogenesis 1
ssa11	LOC106563271	66340506	66382645	zinc finger protein 618-like
ssa12	LOC106566150	83440324	83569524	kinesin heavy chain isoform 5A-like
ssa12	LOC106566244	90131368	90147215	inositol-pentakisphosphate 2-kinase-like
ssa13	LOC106567000	32769414	32851876	disco-interacting protein 2 homolog B-A
ssa13	LOC106568111	84873016	84901978	trace amine-associated receptor 13c-like
ssa13	wwc1	70955397	71005911	WW and C2 domain containing 1
ssa14	LOC106568933	13949128	13959922	choline-phosphate cytidylyltransferase A-like
ssa15	itpk1	40858484	40928332	inositol-tetrakisphosphate 1-kinase
ssa15	LOC106571359	31283077	31391251	protein 4.1-like
ssa15	LOC106571529	36312596	36414587	protein enabled homolog
ssa15	LOC106571754*	51217329	51409612	utrophin-like
ssa15	stxbp5	52022403	52213226	syntaxin binding protein 5 (tomosyn)
ssa15	tmem251	40830323	40857942	transmembrane protein 251
ssa16	bend7	23356567	23364386	BEN domain containing 7
ssa16	LOC106573414	23323009	23327401	kelch repeat and BTB domain-containing protein 13-like
ssa16	LOC106573416"	23343331	23354806	selenide2C water dikinase 1-like
ssa16	LOC106573427	23563538	23623694	C2 domain-containing protein 5-like
ssa16	LOC106573506	27440279	27710925	nuclear receptor ROR-alpha-like
ssa16	LOC106573509	22715465	23002856	SH3 and multiple ankyrin repeat domains protein 3-like
ssa16	LOC106573702	32362952	33047110	cadherin-13-like
ssa17	LOC106576270	44492421	44504981	transcription factor Spi-C-like
ssa18	LOC106576912	12854410	12857878	T-cell leukemia homeobox protein 1-like
ssa20	atp2a2	18887587	18934973	ATPase2C Ca++ transporting2C cardiac muscle2C slow twitch 2
ssa21	baz2b	14640471	14732354	bromodomain adjacent to zinc finger domain2C 2B
ssa21	LOC106582019	24697123	24759127	guanine nucleotide exchange factor DBS-like
ssa22	LOC106582736	9175335	9458601	protein FAM19A2-like
ssa22	LOC106582832	12436441	12558962	SLIT-ROBO Rho GTPase-activating protein 2-like
ssa22	poc1a	59808411	59890914	POC1 centriolar protein A
ssa24	ndufaf2	7520982	7558381	NADH dehydrogenase (ubiquinone) complex I2C assembly factor 2
ssa25	LOC106586888	47217827	47228437	interferon-induced GTP-binding protein Mx-like
ssa25	LOC106586889	47139132	47161992	interferon-induced GTP-binding protein Mx-like
ssa25	LOC106586890	47175785	47193272	interferon-induced GTP-binding protein Mx2-like
ssa26	LOC106587022	1866292	1879289	protein FAM160B2-like
ssa27	LOC106588883	28636513	28673938	inactive phospholipase C-like protein 2
ssa27	**tbc1d5**	28678800	28706730	TBC1 domain family 2C member 5

Table 2. Candidate genes under positive selection based on overlap between the Bayenv2.0 and BayeScan2.1 tests. Genes harboring SNPs that are among 50 the most differentiated SNPs are marked with (*****) if they appear in *Barents & White Seas vs. Ladoga* comparison, and with (") if they are in *Barents & White Seas vs. Onega* comparison. Genes that harbor non-synonymous outlier SNPs are noted in **bold**.